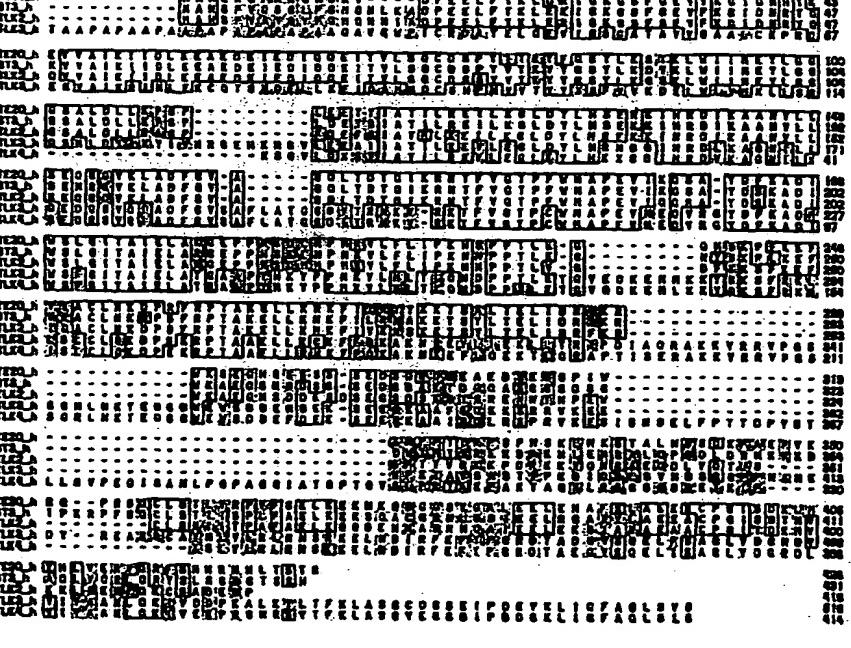




3

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 9/00		A2	(11) International Publication Number: WO 99/53036		
(21) International Application Number: PCT/US99/08150		(43) International Publication Date: 21 October 1999 (21.10.99)			
(22) International Filing Date: 13 April 1999 (13.04.99)					
(30) Priority Data: 60/081,784 14 April 1998 (14.04.98) US					
(71) Applicant (for all designated States except US): SUGEN, INC. [US/US]; 230 East Grand Avenue, South San Francisco, CA 94080 (US).		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).			
(72) Inventors; and		Published			
(75) Inventors/Applicants (for US only): PLOWMAN, Gregory [US/US]; 4 Honeysuckle Lane, San Carlos, CA 94070 (US). MARTINEZ, Ricardo [US/US]; 984 Cartier Lane, Foster City, CA 94404 (US). WHYTE, David [US/US]; 2623 Barclay Way, Belmont, CA 94002 (US).		Without international search report and to be republished upon receipt of that report.			
(74) Agents: WARBURG, Richard, J. et al.; Lyon & Lyon LLP, Suite 4700, 633 West Fifth Street, Los Angeles, CA 90071-2066 (US).					
(54) Title: STE20-RELATED PROTEIN KINASES					
					
(57) Abstract					
<p>The present invention relates to the novel kinase polypeptides STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5, nucleotide sequences encoding the novel kinase polypeptides, as well as various products and methods useful for the diagnosis and treatment of various kinase-related diseases and conditions.</p>					

*** (Referred to in PCT Gazette No. 7/2000, Section II)**

ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5, nucleotide sequences encoding the novel kinase polypeptides, as well as various products and methods useful for the diagnosis and treatment of various kinase-related diseases and conditions.

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

DESCRIPTION

STE20-RELATED PROTEIN KINASES

5

RELATED APPLICATIONS

The present application claims priority to U.S. Provisional Patent Application Serial No. 60/081,784 by Plowman and Martinez, entitled STE20-Related Protein kinases, filed April 14, 1998 (Lyon & Lyon Docket No. 10 232/279), hereby incorporated by reference herein in its entirety, including any drawings, tables, or figures.

FIELD OF THE INVENTION

15 The present invention relates to novel kinase polypeptides, nucleotide sequences encoding the novel kinase polypeptides, as well as various products and methods useful for the diagnosis and treatment of various kinase-related diseases and conditions.

20

BACKGROUND OF THE INVENTION

The following description of the background of the invention is provided to aid in understanding the invention, but is not admitted to be or to describe prior art to the invention.

25

Cellular signal transduction is a fundamental mechanism whereby external stimuli that regulate diverse cellular processes are relayed to the interior of cells. One of the key biochemical mechanisms of signal transduction involves the reversible phosphorylation of proteins, which enables regulation of the activity of mature proteins by altering their structure and function.

30

The best characterized protein kinases in eukaryotes phosphorylate proteins on the hydroxyl moiety of serine, threonine and tyrosine residues. These kinases largely fall

into two groups, those specific for phosphorylating serines and threonines, and those specific for phosphorylating tyrosines. Some kinases, referred to as "dual specificity" kinases, are able to phosphorylate on tyrosine as well as serine/threonine residues.

Protein kinases can also be characterized by their location within the cell. Some kinases are transmembrane receptor-type proteins capable of directly altering their catalytic activity in response to the external environment such as the binding of a ligand. Others are non-receptor-type proteins lacking any transmembrane domain. They can be found in a variety of cellular compartments from the inner surface of the cell membrane to the nucleus.

Many kinases are involved in regulatory cascades wherein their substrates may include other kinases whose activities are regulated by their phosphorylation state. Ultimately the activity of some downstream effector is modulated by phosphorylation resulting from activation of such a pathway.

Protein kinases are one of the largest families of eukaryotic proteins with several hundred known members. These proteins share a 250-300 amino acid domain that can be subdivided into 12 distinct subdomains that comprise the common catalytic core structure. These conserved protein motifs have recently been exploited using PCR-based cloning strategies leading to a significant expansion of the known kinases.

Multiple alignment of the sequences in the catalytic domain of protein kinases and subsequent parsimony analysis permits the segregation of related kinases into distinct branches or subfamilies including: tyrosine kinases, cyclic-nucleotide-dependent kinases, calcium/calmodulin kinases, cyclin-dependent kinases and MAP-kinases, serine-

threonine kinase receptors, and several other less defined subfamilies.

SUMMARY OF THE INVENTION

5 Through the use of a targeted PCR cloning strategy and of a "motif extraction" bioinformatics script, mammalian members of the STE20-kinase family have been identified as part of the present invention. Multiple alignment and parsimony analysis of the catalytic domain of all of these
10 STE20-family members reveals that these proteins cluster into 9 distinct subgroups. Classification in this manner has proven highly accurate not only in predicting motifs present in the remaining non-catalytic portion of each protein, but also in their regulation, substrates, and
15 signaling pathways. The present invention includes the partial or complete sequence of new members of the STE20-family, their classification, predicted or deduced protein structure, and a strategy for elucidating their biologic and therapeutic relevance.

20 Thus, a first aspect of the invention features an isolated, enriched, or purified nucleic acid molecule encoding a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5.

25 By "isolated" in reference to nucleic acid is meant a polymer of nucleotides conjugated to each other, including DNA and RNA, that is isolated from a natural source or that is synthesized. The isolated nucleic acid of the present invention is unique in the sense that it is not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal cellular (i.e., chromosomal) environment. Thus, the sequence may be in a cell-free

solution or placed in a different cellular environment. The term does not imply that the sequence is the only nucleotide chain present, but that it is essentially free (about 90 - 95% pure at least) of non-nucleotide material naturally associated with it, and thus is distinguished from isolated chromosomes.

5 By the use of the term "enriched" in reference to nucleic acid is meant that the specific DNA or RNA sequence constitutes a significantly higher fraction (2 - 5 fold) of the total DNA or RNA present in the cells or solution of interest than in normal or diseased cells or in the cells from which the sequence was taken. This could be caused by 10 a person by preferential reduction in the amount of other DNA or RNA present, or by a preferential increase in the amount of the specific DNA or RNA sequence, or by a combination of the two. However, it should be noted that 15 enriched does not imply that there are no other DNA or RNA sequences present, just that the relative amount of the sequence of interest has been significantly increased. The term "significant" is used to indicate that the level of 20 increase is useful to the person making such an increase, and generally means an increase relative to other nucleic acids of about at least 2 fold, more preferably at least 5 to 10 fold or even more. The term also does not imply that 25 there is no DNA or RNA from other sources. The other source DNA may, for example, comprise DNA from a yeast or bacterial genome, or a cloning vector such as pUC19. This term distinguishes from naturally occurring events, such as viral 30 infection, or tumor type growths, in which the level of one mRNA may be naturally increased relative to other species of mRNA. That is, the term is meant to cover only those situations in which a person has intervened to elevate the proportion of the desired nucleic acid.

It is also advantageous for some purposes that a nucleotide sequence be in purified form. The term "purified" in reference to nucleic acid does not require absolute purity (such as a homogeneous preparation).
5 Instead, it represents an indication that the sequence is relatively more pure than in the natural environment (compared to the natural level this level should be at least 2-5 fold greater, e.g., in terms of mg/mL). Individual clones isolated from a cDNA library may be purified to 10 electrophoretic homogeneity. The claimed DNA molecules obtained from these clones could be obtained directly from total DNA or from total RNA. The cDNA clones are not naturally occurring, but rather are preferably obtained via manipulation of a partially purified naturally occurring 15 substance (messenger RNA). The construction of a cDNA library from mRNA involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection of the cells carrying the cDNA library. Thus, the process 20 which includes the construction of a cDNA library from mRNA and isolation of distinct cDNA clones yields an approximately 10⁶-fold purification of the native message. Thus, purification of at least one order of magnitude, preferably two or three orders, and more preferably four or 25 five orders of magnitude is expressly contemplated.

By a "kinase polypeptide" is meant 32 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7, or the corresponding full-length amino acid sequence; 250 (preferably 255, more preferably 260, most preferably 270) or more contiguous amino acids set forth in the amino acid sequence SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, or SEQ ID NO:105, or the

corresponding full-length amino acid sequence; 27
(preferably 30, more preferably 40, most preferably 45) or
more contiguous amino acids set forth in the amino acid
sequence SEQ ID NO:18; 16 (preferably 20, more preferably
5 25, most preferably 35) or more contiguous amino acids set
forth in the amino acid sequence SEQ ID NO:22, SEQ ID NO:23,
SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, or SEQ ID NO:103
or the corresponding full-length amino acid sequence; 6
(preferably 10, more preferably 15, most preferably 25) or
10 more contiguous amino acids set forth in the amino acid
sequence of SEQ ID NO:97 or SEQ ID NO:99, 22 (preferably 30,
more preferably 35, most preferably 45) or more contiguous
amino acids set forth in the amino acid sequence of SEQ ID
NO:101, or the corresponding full-length amino acid
15 sequence; 78 (preferably 80, more preferably 85, most
preferably 90) or more contiguous amino acids set forth in
the amino acid sequence SEQ ID NO:107 or functional
derivatives thereof as described herein. For sequences for
which the full-length sequence is not given, the remaining
20 sequences can be determined using methods well-known to
those in the art and are intended to be included in the
invention. In certain aspects, polypeptides of 100, 200,
300 or more amino acids are preferred. The kinase
25 polypeptide can be encoded by a full-length nucleic acid
sequence or any portion of the full-length nucleic acid
sequence, so long as a functional activity of the
polypeptide is retained, not to include fragments containing
only amino acids 1-22 of SEQ ID NO:13 or only amino acids 1-
33 of SEQ ID NO:107.

30 The amino acid sequence will be substantially similar
to the sequence shown in SEQ ID NO:5, SEQ ID NO:6, SEQ ID
NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID
NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID

NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID
NO:100, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, or
the corresponding full-length amino acid sequence, or
fragments thereof, not to include fragments consisting only
5 of the amino acid sequences 1-22 of SEQ ID NO:13 or 1-33 of
SEQ ID NO:107. A sequence that is substantially similar to
the sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ
ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID
NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID
10 NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID
NO:103, SEQ ID NO:105, or SEQ ID NO:107 will preferably have
at least 90% identity (more preferably at least 95% and most
preferably 99-100%) to the sequence of SEQ ID NO:5, SEQ ID
15 NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15,
SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ
ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID
NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107.

By "identity" is meant a property of sequences that
measures their similarity or relationship. Identity is
20 measured by dividing the number of identical residues by the
total number of residues and gaps and multiplying the
product by 100. "Gaps" are spaces in an alignment that are
the result of additions or deletions of amino acids. Thus,
two copies of exactly the same sequence have 100% identity,
25 but sequences that are less highly conserved, and have
deletions, additions, or replacements, may have a lower
degree of identity. Those skilled in the art will recognize
that several computer programs are available for determining
sequence identity using standard parameters, for example
Blast (Altschul, et al. (1997) Nucleic Acids Res. 25:3389-
3402), Blast2 (Altschul, et al. (1990) J. mol. biol.
30 215:403-410), and Smith-Waterman (Smith, et al. (1981) J.
Mol. Biol. 147:195-197).

In preferred embodiments, the invention features isolated, enriched, or purified nucleic acid molecules encoding a kinase polypeptide comprising a nucleotide sequence that: (a) encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107; (b) is the complement of the nucleotide sequence of (a); (c) hybridizes under highly stringent conditions to the nucleotide molecule of (a) and encodes a naturally occurring kinase polypeptide; (d) encodes a kinase polypeptide having the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, except that it lacks one or more, but not all, of the following segments of amino acid residues: 1-21, 22-274, or 275-416 of SEQ ID NO:5, 1-31, 32-308, 309-489 or 490-516 of SEQ ID NO:6, 1-178 or 179-414 of SEQ ID NO:7, 1-22, 23-289, 290-526, 527-640, 641-896, or 897-1239 of SEQ ID NO:13, 1-255, 256-442, 443-626, 627-954, or 955-1297 of SEQ ID NO:14, 1-255, 256-476, 477-680, 681-983, or 984-1326 of SEQ ID NO:15, 1-13, 14-273, 274-346, 347-534, or 535-894 of SEQ ID NO:18, 1-21, 22-277, 278-427, 428-637, 638-751, or 752-898 of SEQ ID NO:22, 1-66, 67-215, 216-425, 426-539, 540-786, or 787-887 of SEQ ID NO:23, 1-25, 26-273, 274-422, 423-632, or 633-748 of SEQ ID NO:24, 1-51, 52-224, 225-393, 394-658, or 659-681 of SEQ ID NO:29, 1-25, 26-281, 284-430, 431-640, 641-754, 755-901, or 902-1001 of SEQ ID NO:31, 1-10, 11-321, or 322-373 of SEQ ID NO:97, 1-57, 58-369, or 370-418 of SEQ ID NO:99, 1-52, 53-173, 174-

307, 308-572, or 573-591 of SEQ ID NO:103, 1-24, 25-289,
290-397, 398-628, 629-872, or 873-1227 of SEQ ID NO:105, or
1-33, 34-294, 295-337, 338-472, 473-724, or 725-968 of SEQ
ID NO:107; (e) is the complement of the nucleotide sequence
5 of (d); (f) encodes a polypeptide having the amino acid
sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7,
SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ
ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID
NO:31; SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:103, SEQ ID
10 NO:105, or SEQ ID NO:107 from amino acid residues 1-21, 22-
274, or 275-416 of SEQ ID NO:5, 1-31, 32-308, 309-489, or
490-516 of SEQ ID NO:6, 1-178 or 179-414 of SEQ ID NO:7, 23-
289, 290-526, 527-640, 641-896, or 897-1239 of SEQ ID NO:13,
1-255, 256-442, 443-626, 627-954, or 955-1297 of SEQ ID
15 NO:14, 1-255, 256-476, 477-680, 681-983, or 984-1326 of SEQ
ID NO:15, 1-13, 14-273, 274-346, 347-534, or 535-894 of SEQ
ID NO:18, 1-21, 22-277, 278-427, 428-637, 638-751, or 752-
898 of SEQ ID NO:22, 1-66, 67-215, 216-425, 426-539, 540-
786, or 787-887 of SEQ ID NO:23, 1-25, 26-273, 274-422, 423-
20 632, or 633-748 of SEQ ID NO:24, 1-51, 52-224, 225-393, 394-
658, or 659-681 of SEQ ID NO:29, 1-25, 26-281, 282-430, 431-
640, 641-754, 755-901, or 902-1001 of SEQ ID NO:31, 1-10,
11-321, or 322-373 of SEQ ID NO:97, 1-57, 58-369, or 370-418
of SEQ ID NO:99, 1-52, 53-173, 174-307, 308-572, or 573-591
25 of SEQ ID NO:103, 1-24, 25-289, 290-397, 398-628, 629-872,
or 873-1227 of SEQ ID NO:105, or 1-33, 34-294, 295-337, 338-
472, 473-724, or 725-968 of SEQ ID NO:107; (g) is the
complement of the nucleotide sequence of (f); (h) encodes a
polypeptide having the amino acid sequence set forth in SEQ
30 ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID
NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID
NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID
NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID

NO:105, or SEQ ID NO:107, except that it lacks one or more of the domains selected from the group consisting of a N-terminal domain, a catalytic domain, a C-terminal domain, a coiled-coil structure region, a proline-rich region, a 5 spacer region, an insert, and a C-terminal tail; or (i) is the complement of the nucleotide sequence of (h).

The term "complement" refers to two nucleotides that can form multiple favorable interactions with one another. For example, adenine is complementary to thymine as they can 10 form two hydrogen bonds. Similarly, guanine and cytosine are complementary since they can form three hydrogen bonds.

A nucleotide sequence is the complement of another nucleotide sequence if all of the nucleotides of the first sequence are complementary to all of the nucleotides of the 15 second sequence.

The term "domain" refers to a region of a polypeptide which contains a particular function. For instance, N-terminal or C-terminal domains of signal transduction proteins can serve functions including, but not limited to, 20 binding molecules that localize the signal transduction molecule to different regions of the cell or binding other signaling molecules directly responsible for propagating a particular cellular signal. Some domains can be expressed separately from the rest of the protein and function by themselves, while others must remain part of the intact 25 protein to retain function. The latter are termed functional regions of proteins and also relate to domains.

The term "N-terminal domain" refers to the extracatalytic region located between the initiator methionine and the catalytic domain of the protein kinase. 30 The N-terminal domain can be identified following a Smith-Waterman alignment of the protein sequence against the non-redundant protein database to define the N-terminal boundary.

of the catalytic domain. Depending on its length, the N-terminal domain may or may not play a regulatory role in kinase function. An example of a protein kinase whose N-terminal domain has been shown to play a regulatory role is PAK65, which contains a CRIB motif used for Cdc42 and rac binding (Burbelo, P.D. et al. (1995) J. Biol. Chem. 270, 29071-290740).

The N-terminal domain spans amino acid residues 1-21 of the sequence set forth in SEQ ID NO:5, amino acid residues 1-31 of the sequence set forth in SEQ ID NO:6, amino acid residues 1-22 of the sequence set forth in SEQ ID NO:13, amino acid residues 1-13 of the sequence set forth in SEQ ID NO:18, amino acid residues 1-21 of the sequence set forth in SEQ ID NO:22, amino acid residues 1-25 of the sequence set forth in SEQ ID NO:24, amino acid residues 1-51 of the sequence set forth in SEQ ID NO:29, amino acid residues 1-25 of the sequence set forth in SEQ ID NO:31, amino acid residues 1-57 of the sequence set forth in SEQ ID NO:99, amino acid residues 1-52 of the sequence set forth in SEQ ID NO:103, amino acid residues 1-24 of the sequence set forth in SEQ ID NO:105, or amino acid residues 1-33 of the sequence set forth in SEQ ID NO:107.

The term "catalytic domain" refers to a region of the protein kinase that is typically 25-300 amino acids long and is responsible for carrying out the phosphate transfer reaction from a high-energy phosphate donor molecule such as ATP or GTP to itself (autophosphorylation) or to other proteins (exogenous phosphorylation). The catalytic domain of protein kinases is made up of 12 subdomains that contain highly conserved amino acid residues, and are responsible for proper polypeptide folding and for catalysis. The catalytic domain can be identified following a Smith-

Waterman alignment of the protein sequence against the non-redundant protein database.

The catalytic domain spans amino acid residues 22-274 of the sequence set forth in SEQ ID NO:5, residues 32-308 of the sequence set forth in SEQ ID NO:6, residues 1-178 of the sequence set forth in SEQ ID NO:7, residues 23-289 of the sequence set forth in SEQ ID NO:13, residues 1-255 of the sequence set forth in SEQ ID NO:14, residues 1-255 of the sequence set forth in SEQ ID NO:15, residues 14-273 of the sequence set forth in SEQ ID NO:18, residues 22-277 of the sequence set forth in SEQ ID NO:22, residues 1-66 of the sequence set forth in SEQ ID NO:23, residues 26-273 of the sequence set forth in SEQ ID NO:24, residues 394-658 of the sequence set forth in SEQ ID NO:29, residues 26-281 of the sequence set forth in SEQ ID NO:31, residues 1-278 of the sequence set forth in SEQ ID NO:97, residues 58-369 of the sequence set forth in SEQ ID NO:99, residues 1-103 of the sequence set forth in SEQ ID NO:101, residues 308-572 of the sequence set forth in SEQ ID NO:103, residues 25-289 of the sequence set forth in SEQ ID NO:105, or residues 34-294 of the sequence set forth in SEQ ID NO:107.

The term "catalytic activity", as used herein, defines the rate at which a kinase catalytic domain phosphorylates a substrate. Catalytic activity can be measured, for example, by determining the amount of a substrate converted to a phosphorylated product as a function of time. Catalytic activity can be measured by methods of the invention by holding time constant and determining the concentration of a phosphorylated substrate after a fixed period of time. Phosphorylation of a substrate occurs at the active-site of a protein kinase. The active-site is normally a cavity in which the substrate binds to the protein kinase and is phosphorylated.

The term "substrate" as used herein refers to a molecule phosphorylated by a kinase of the invention. Kinases phosphorylate substrates on serine/threonine or tyrosine amino acids. The molecule may be another protein or a polypeptide.

5 The term "C-terminal domain" refers to the region located between the catalytic domain or the last (located closest to the C-terminus) functional domain and the carboxy-terminal amino acid residue of the protein kinase.

10 By "functional" domain is meant any region of the polypeptide that may play a regulatory or catalytic role as predicted from amino acid sequence homology to other proteins or by the presence of amino acid sequences that may give rise to specific structural conformations (i.e. coiled-coils). The C-terminal domain can be identified by using a Smith-Waterman alignment of the protein sequence against the non-redundant protein database to define the C-terminal boundary of the catalytic domain or of any functional C-terminal extracatalytic domain. Depending on its length and amino acid composition, the C-terminal domain may or may not play a regulatory role in kinase function. An example of a protein kinase whose C-terminal domain may play a regulatory role is PAK3 which contains a heterotrimeric G_b subunit-binding site near its C-terminus (Leeuw, T. et al (1998) Nature, 391, 191-195).

15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
The C-terminal domain spans amino acid residues 275-416 of the sequence set forth in SEQ ID NO:5, residues 309-489 of the sequence set forth in SEQ ID NO:6, residues 179-414 of the sequence set forth in SEQ ID NO:7, residues 897-1239 of the sequence set forth in SEQ ID NO:13, residues 955-1297 of the sequence set forth in SEQ ID NO:14, residues 984-1326 of the sequence set forth in SEQ ID NO:15, residues 535-894 of the sequence set forth in SEQ ID NO:18, residues 752-898

of the sequence set forth in SEQ ID NO:22, residues 279-330
of the sequence set forth in SEQ ID NO:97, residues 370-418
of the sequence set forth in SEQ ID NO:99, or residues 873-
1227 of the sequence set forth in SEQ ID NO:105.

5 The term "signal transduction pathway" refers to the
molecules that propagate an extracellular signal through the
cell membrane to become an intracellular signal. This
signal can then stimulate a cellular response. The
10 polypeptide molecules involved in signal transduction
processes are typically receptor and non-receptor protein
tyrosine kinases, receptor and non-receptor protein
phosphatases, SRC homology 2 and 3 domains, phosphotyrosine
binding proteins (SRC homology 2 (SH2) and phosphotyrosine
binding (PTB and PH) domain containing proteins), proline-
15 rich binding proteins (SH3 domain containing proteins),
nucleotide exchange factors, and transcription factors.

The term "coiled-coil structure region" as used herein,
refers to a polypeptide sequence that has a high probability
of adopting a coiled-coil structure as predicted by computer
20 algorithms such as COILS (Lupas, A. (1996) Meth. Enzymology
266:513-525). Coiled-coils are formed by two or three
amphipathic α -helices in parallel. Coiled-coils can bind to
coiled-coil domains of other polypeptides resulting in homo-
or heterodimers (Lupas, A. (1991) Science 252:1162-1164).
25 Coiled-coil-dependent oligomerization has been shown to be
necessary for protein function including catalytic activity
of serine/threonine kinases (Roe, J. et al. (1997) J. Biol.
Chem. 272:5838-5845).

30 The coiled-coil structure region spans amino acid
residues 290-526 of the sequence set forth in SEQ ID NO:13,
residues 256-442 of the sequence set forth in SEQ ID NO:14,
residues 256-476 of the sequence set forth in SEQ ID NO:15,
residues 428-637 of the sequence set forth in SEQ ID NO:22,

residues 216-425 or 540-786 of the sequence set forth in SEQ ID NO:23, residues 423-632 of the sequence set forth in SEQ ID NO:24, residues 431-640 or 755-901 of the sequence set forth in SEQ ID NO:31, residues 291-398 or 629-668 of the sequence set forth in SEQ ID NO:105, or residues 473-724 or 725-968 of the sequence set forth in SEQ ID NO:107.

The term "proline-rich region" as used herein, refers to a region of a protein kinase whose proline content over a given amino acid length is higher than the average content of this amino acid found in proteins(i.e., >10%). Proline-rich regions are easily discernable by visual inspection of amino acid sequences and quantitated by standard computer sequence analysis programs such as the DNASTar program EditSeq. Proline-rich regions have been demonstrated to participate in regulatory protein -protein interactions. Among these interactions, those that are most relevant to this invention involve the "PxxP" proline rich motif found in certain protein kinases (i.e., human PAK1) and the SH3 domain of the adaptor molecule Nck (Galisteo, M.L. et al. (1996) J. Biol. Chem. 271:20997-21000). Other regulatory interactions involving "PxxP" proline-rich motifs include the WW domain (Sudol, M. (1996) Prog. Biochys. Mol. Bio. 65:113-132).

The proline-rich region spans amino acid residues 527-640 of the sequence set forth in SEQ ID NO:13, residues 443-626 of the sequence set forth in SEQ ID NO:14, residues 477-680 of the sequence set forth in SEQ ID NO:15, residues 347-534 of the sequence set forth in SEQ ID NO:18, residues 398-628 of the sequence set forth in SEQ ID NO:105, or residues 338-472 of the sequence set forth in SEQ ID NO:107.

The term "spacer region" as used herein, refers to a region of the protein kinase located between predicted functional domains. The spacer region has no detectable

homology to any amino acid sequence in the database, and can be identified by using a Smith-Waterman alignment of the protein sequence against the non-redundant protein database to define the C- and N-terminal boundaries of the flanking functional domains. Spacer regions may or may not play a fundamental role in protein kinase function. Precedence for the regulatory role of spacer regions in kinase function is provided by the role of the src kinase spacer in inter-domain interactions (Xu, W. et al. (1997) Nature 385:595-602).

The spacer region spans amino acid residues 641-896 of the sequence set forth in SEQ ID NO:13, residues 627-954 of the sequence set forth in SEQ ID NO:14, residues 681-983 of the sequence set forth in SEQ ID NO:15, residues 274-346 of the sequence set forth in SEQ ID NO:18, residues 278-427 or 638-751 of the sequence set forth in SEQ ID NO:22, residues 67-215 or 426-539 of the sequence set forth in SEQ ID NO:23, residues 274-422 or 633-748 of the sequence set forth in SEQ ID NO:24, residues 225-393 of the sequence set forth in SEQ ID NO:29, residues 282-430 or 641-754 of the sequence set forth in SEQ ID NO:31, residues 174-307 of the sequence set forth in SEQ ID NO:103, residues 669-872 of the sequence set forth in SEQ ID NO:105, or residues 295-337 of the sequence set forth in SEQ ID NO:107.

The term "insert" as used herein refers to a portion of a protein kinase that is absent from a close homolog. Inserts may or may not be the product alternative splicing of exons. Inserts can be identified by using a Smith-Waterman sequence alignment of the protein sequence against the non-redundant protein database, or by means of a multiple sequence alignment of homologous sequences using the DNASTar program Megalign. Inserts may play a functional role by presenting a new interface for protein-protein

interactions, or by interfering with such interactions. Inserts span amino acid residues 52-224 of the sequence set forth in SEQ ID NO:29 or residues 53-173 of the sequence set forth in SEQ ID NO:103.

5 The term "C-terminal tail" as used herein, refers to a C-terminal domain of a protein kinase, that by homology extends or protrudes past the C-terminal amino acid of its closest homolog. C-terminal tails can be identified by using a Smith-Waterman sequence alignment of the protein sequence against the non-redundant protein database, or by means of a multiple sequence alignment of homologous sequences using the DNASTar program Megalign. Depending on its length, a C-terminal tail may or may not play a regulatory role in kinase function.

10 The C-terminal tail spans amino acid residues 490-516 of the sequence set forth in SEQ ID NO:6, residues 787-887 of the sequence set forth in SEQ ID NO:23, residues 659-681 of the sequence set forth in SEQ ID NO:29, residues 994-1093 of the sequence set forth in SEQ ID NO:31, or residues 573-591 of the sequence set forth in SEQ ID NO:103.

15 Various low or high stringency hybridization conditions may be used depending upon the specificity and selectivity desired. These conditions are well-known to those skilled in the art. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 20 contiguous nucleotides, more preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 50 contiguous nucleotides, most preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 100 contiguous nucleotides. In some instances, the conditions

may prevent hybridization of nucleic acids having more than 5 mismatches in the full-length sequence.

By stringent hybridization assay conditions is meant hybridization assay conditions at least as stringent as the following: hybridization in 50% formamide, 5X SSC, 50 mM NaH₂PO₄, pH 6.8, 0.5% SDS, 0.1 mg/mL sonicated salmon sperm DNA, and 5X Denhart solution at 42 °C overnight; washing with 2X SSC, 0.1% SDS at 45 °C; and washing with 0.2X SSC, 0.1% SDS at 45 °C. Under some of the most stringent hybridization assay conditions, the second wash can be done with 0.1X SSC at a temperature up to 70 °C (Berger et al. (1987) Guide to Molecular Cloning Techniques pg 421, hereby incorporated by reference herein including any figures, tables, or drawings.). However, other applications may require the use of conditions falling between these sets of conditions.

Methods of determining the conditions required to achieve desired hybridizations are well-known to those with ordinary skill in the art, and are based on several factors, including but not limited to, the sequences to be hybridized and the samples to be tested.

In other preferred embodiments, the invention features isolated, enriched, or purified nucleic acid molecules encoding kinase polypeptides, further comprising a vector or promoter effective to initiate transcription in a host cell.

The invention also features recombinant nucleic acid, preferably in a cell or an organism. The recombinant nucleic acid may contain a sequence set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:27, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, or SEQ ID NO:106, or a functional derivative thereof and a vector or a promoter effective to initiate transcription in a host cell. The

recombinant nucleic acid can alternatively contain a transcriptional initiation region functional in a cell, a sequence complementary to an RNA sequence encoding a kinase polypeptide and a transcriptional termination region functional in a cell. Specific vectors and host cell combinations are discussed herein.

The term "vector" relates to a single or double-stranded circular nucleic acid molecule that can be transfected into cells and replicated within or independently of a cell genome. A circular double-stranded nucleic acid molecule can be cut and thereby linearized upon treatment with restriction enzymes. An assortment of nucleic acid vectors, restriction enzymes, and the knowledge of the nucleotide sequences cut by restriction enzymes are readily available to those skilled in the art. A nucleic acid molecule encoding a kinase can be inserted into a vector by cutting the vector with restriction enzymes and ligating the two pieces together.

The term "transfected" defines a number of methods to insert a nucleic acid vector or other nucleic acid molecules into a cellular organism. These methods involve a variety of techniques, such as treating the cells with high concentrations of salt, an electric field, detergent, or DMSO to render the outer membrane or wall of the cells permeable to nucleic acid molecules of interest or use of various viral transduction strategies.

The term "promoter" as used herein, refers to nucleic acid sequence needed for gene sequence expression. Promoter regions vary from organism to organism, but are well known to persons skilled in the art for different organisms. For example, in prokaryotes, the promoter region contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when

transcribed into RNA, will signal synthesis initiation. Such regions will normally include those 5'-non-coding sequences involved with initiation of transcription and translation, such as the TATA box, capping sequence, CAAT sequence, and the like.

In preferred embodiments, the isolated nucleic acid comprises, consists essentially of, or consists of a nucleic acid sequence set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:27, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100 SEQ ID NO:102, SEQ ID NO:104, or SEQ ID NO:106, or the corresponding full-length sequence, encodes the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, or the corresponding full-length amino acid sequence, a functional derivative thereof, or at least 40, 45, 50, 60, 100, 200, or 300 contiguous amino acids of SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7; or of the corresponding full-length amino acid sequence; at least 250, 255, 275, 300, or 400 contiguous amino acids of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, or of the corresponding full-length amino acid sequence; at least 27, 30, 35, 40, 50, 100, 200, or 300 contiguous amino acids of SEQ ID NO:18; at least 16, 25, 35, 50, 100, 200, or 300 contiguous amino acids of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, or SEQ ID NO:103, or of the corresponding full-length amino acid sequence; 6 (preferably 10, more preferably 15, most preferably 25) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:97 or SEQ ID NO:99, or the corresponding full-

length amino acid sequence; 22 (preferably 30, more preferably 35, most preferably 45) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:101, or the corresponding full-length amino acid sequence; or at least 80, 85, 90, 100, 200, or 300 contiguous amino acids of SEQ ID NO:107, or functional derivatives thereof. The kinase polypeptides, selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5, comprise, consist essentially of, or consist of at least at least 40, 45, 50, 60, 100, 200, or 300 contiguous amino acids of SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7; at least 250, 255, 275, 300, or 400 contiguous amino acids of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, or SEQ ID NO:105; at least 27, 30, 35, 40, 50, 100, 200, or 300 contiguous amino acids of SEQ ID NO:18; at least 35, 40, 45, 50, 100, 200, or 300 contiguous amino acids of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31 or SEQ ID NO:103; 6 (preferably 10, more preferably 15, most preferably 25) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:97 or SEQ ID NO:99; 22 (preferably 30, more preferably 35, most preferably 45) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:101; or at least 80, 85, 90, 100, 200, or 300 contiguous amino acids of SEQ ID NO:107, or the corresponding full-length sequences or derivatives thereof. The nucleic acid may be isolated from a natural source by cDNA cloning or by subtractive hybridization. The natural source may be mammalian, preferably human, blood, semen, or tissue, and the nucleic acid may be synthesized by the triester method or by using an automated DNA synthesizer.

The term "mammal" refers preferably to such organisms as mice, rats, rabbits, guinea pigs, sheep, and goats, more

preferably to cats, dogs, monkeys, and apes, and most preferably to humans.

In yet other preferred embodiments, the nucleic acid is a conserved or unique region, for example those useful for:
5 the design of hybridization probes to facilitate identification and cloning of additional polypeptides, the design of PCR probes to facilitate cloning of additional polypeptides, obtaining antibodies to polypeptide regions, and designing antisense oligonucleotides.

By "conserved nucleic acid regions", are meant regions present on two or more nucleic acids encoding a kinase polypeptide, to which a particular nucleic acid sequence can hybridize under lower stringency conditions. Examples of lower stringency conditions suitable for screening for 10 nucleic acid encoding kinase polypeptides are provided in Abe, et al. (J. Biol. Chem. 19:13361-13368, 1992), hereby incorporated by reference herein in its entirety, including any drawings, figures, or tables. Preferably, conserved regions differ by no more than 5 out of 20 nucleotides, even 15 more preferably 2 out of 20 nucleotides or most preferably 1 out of 20 nucleotides.
20

By "unique nucleic acid region" is meant a sequence present in a nucleic acid coding for a kinase polypeptide that is not present in a sequence coding for any other naturally occurring polypeptide. Such regions preferably 25 encode 32 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7, or the corresponding full-length amino acid sequence; 250 (preferably 255, more preferably 260, most preferably 270) or more contiguous amino acids set forth in 30 the amino acid sequence SEQ ID NO:13, SEQ ID NO:14, or SEQ ID NO:15, or SEQ ID NO:105, or the corresponding full-length

amino acid sequence; 27 (preferably 30, more preferably 40, most preferably 45) or more contiguous amino acids set forth in the amino acid sequence SEQ ID NO:18; 16 (preferably 20, more preferably 25, most preferably 35) or more contiguous 5 amino acids set forth in the amino acid sequence SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, or SEQ ID NO:103, or the corresponding full-length amino acid sequence; 6 (preferably 10, more preferably 15, most preferably 25) or more contiguous amino acids set forth 10 in the amino acid sequence of SEQ ID NO:97 or SEQ ID NO:99, 22 (preferably 30, more preferably 35, most preferably 45) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:101, or the corresponding full-length 15 amino acid sequence; or 78 (preferably 80, more preferably 85, most preferably 90) or more contiguous amino acids set forth in the amino acid sequence SEQ ID NO:107, or functional derivatives thereof. In particular, a unique 20 nucleic acid region is preferably of mammalian origin.

A second aspect of the invention features a nucleic 20 acid probe for the detection of nucleic acid encoding a kinase polypeptide in a sample, wherein said polypeptide is selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5. Preferably, the nucleic acid probe 25 encodes a kinase polypeptide that is a fragment of the protein encoded by the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, or the corresponding full-length 30 amino acid sequences, not to include fragments consisting only of amino acids 1-22 of SEQ ID NO:13 or amino acids 1-33

of SEQ ID NO:107. The nucleic acid probe contains a nucleotide base sequence that will hybridize to a sequence set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:27, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, or SEQ ID NO:106, or the corresponding full-length sequence, or a functional derivative thereof.

In preferred embodiments, the nucleic acid probe hybridizes to nucleic acid encoding at least 6, 12, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31 SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, or the corresponding full-length amino acid sequence, or functional derivatives thereof.

Methods for using the probes include detecting the presence or amount of kinase RNA in a sample by contacting the sample with a nucleic acid probe under conditions such that hybridization occurs and detecting the presence or amount of the probe bound to kinase RNA. The nucleic acid duplex formed between the probe and a nucleic acid sequence coding for a kinase polypeptide may be used in the identification of the sequence of the nucleic acid detected (Nelson et al., in Nonisotopic DNA Probe Techniques, Academic Press, San Diego, Kricka, ed., p. 275, 1992, hereby incorporated by reference herein in its entirety, including any drawings, figures, or tables). Kits for performing such methods may be constructed to include a container means having disposed therein a nucleic acid probe.

In a third aspect, the invention describes a recombinant cell or tissue comprising a nucleic acid molecule encoding a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, 5 STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5. In such cells, the nucleic acid may be under the control of the genomic regulatory elements, or may be under the control of exogenous regulatory elements including an exogenous promoter. By "exogenous" it is meant a promoter 10 that is not normally coupled *in vivo* transcriptionally to the coding sequence for the kinase polypeptides.

The polypeptide is preferably a fragment of the protein encoded by the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID 15 NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, or the corresponding full-length amino acid sequence, not to include fragments consisting only of amino acids 1-22 of SEQ ID NO:13 or amino acids 1-33 of SEQ ID 20 NO:107. By "fragment," is meant an amino acid sequence present in a kinase polypeptide. Preferably, such a sequence comprises at least 32, 45, 50, 60, 100, 200, or 300 contiguous amino acids of SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7, or of the corresponding full-length amino acid sequence; at least 250, 255, 275, 300, or 400 contiguous amino acids of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, OR 25 SEQ ID NO:105, or of the corresponding full-length amino acid sequence; at least 27, 30, 35, 40, 50, 100, 200, or 300 contiguous amino acids of SEQ ID NO:18; at least 16, 25, 35, 30 50, 100, 200, or 300 contiguous amino acids of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31 or SEQ ID NO:103, or of the corresponding full-length amino

acid sequence; 6 (preferably 10, more preferably 15, most preferably 25) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:97 or SEQ ID NO:99, 22 (preferably 30, more preferably 35, most preferably 45) or 5 more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:101; at least 78, 85, 90, 100, 200, or 300 contiguous amino acids of SEQ ID NO:107, or the corresponding full-length amino acid sequence; or a functional derivative thereof.

10 In a fourth aspect, the invention features an isolated, enriched, or purified kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5.

15 By "isolated" in reference to a polypeptide is meant a polymer of amino acids (2 or more amino acids) conjugated to each other, including polypeptides that are isolated from a natural source or that are synthesized. The isolated polypeptides of the present invention are unique in the 20 sense that they are not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal cellular environment. Thus, the sequence may be in a cell-free solution or placed in a different cellular 25 environment. The term does not imply that the sequence is the only amino acid chain present, but that it is essentially free (about 90 - 95% pure at least) of non-amino acid material naturally associated with it.

30 By the use of the term "enriched" in reference to a polypeptide is meant that the specific amino acid sequence constitutes a significantly higher fraction (2 - 5 fold) of the total amino acid sequences present in the cells or solution of interest than in normal or diseased cells or in

the cells from which the sequence was taken. This could be caused by a person by preferential reduction in the amount of other amino acid sequences present, or by a preferential increase in the amount of the specific amino acid sequence of interest, or by a combination of the two. However, it should be noted that enriched does not imply that there are no other amino acid sequences present, just that the relative amount of the sequence of interest has been significantly increased. The term significant here is used to indicate that the level of increase is useful to the person making such an increase, and generally means an increase relative to other amino acid sequences of about at least 2-fold, more preferably at least 5- to 10-fold or even more. The term also does not imply that there is no amino acid sequence from other sources. The other source of amino acid sequences may, for example, comprise amino acid sequence encoded by a yeast or bacterial genome, or a cloning vector such as pUC19. The term is meant to cover only those situations in which man has intervened to increase the proportion of the desired amino acid sequence.

It is also advantageous for some purposes that an amino acid sequence be in purified form. The term "purified" in reference to a polypeptide does not require absolute purity (such as a homogeneous preparation); instead, it represents an indication that the sequence is relatively purer than in the natural environment. Compared to the natural level this level should be at least 2-5 fold greater (e.g., in terms of mg/mL). Purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. The substance is preferably free of contamination at a functionally significant level, for example 90%, 95%, or 99% pure.

In preferred embodiments, the kinase polypeptide is a fragment of the protein encoded by the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, or the corresponding full-length amino acid sequences, not to include fragments consisting only of amino acids 1-22 of SEQ ID NO:13 or amino acids 1-33 of SEQ ID NO:107. Preferably, the kinase polypeptide contains at least 32, 45, 50, 60, 100, 200, or 300 contiguous amino acids of SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7, or the corresponding full-length amino acid sequence; at least 250, 255, 275, 300, or 400 contiguous amino acids of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, or SEQ ID NO:105, or the corresponding full-length amino acid sequence; at least 27, 30, 35, 40, 50, 100, 200, or 300 contiguous amino acids of SEQ ID NO:18; at least 16, 25, 35, 50, 100, 200, or 300 contiguous amino acids of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, or SEQ ID NO:103, or the corresponding full-length amino acid sequence; 6 (preferably 10, more preferably 15, most preferably 25) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:97 or SEQ ID NO:99, 22 (preferably 30, more preferably 35, most preferably 45) or more contiguous amino acids set forth in the amino acid sequence of SEQ ID NO:101, or the corresponding full-length amino acid sequence; or at least 78, 85, 90, 100, 200, or 300 contiguous amino acids of SEQ ID NO:107, or a functional derivative thereof.

In preferred embodiments, the kinase polypeptide comprises an amino acid sequence having (a) the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7,

SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107; (b) the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, except that it lacks one or more, but not all, of the following segments of amino acid residues: 1-21, 22-274, or 275-416 of SEQ ID NO:5, 1-31, 32-308, 309-489 or 490-516 of SEQ ID NO:6, 1-178 or 179-414 of SEQ ID NO:7, 1-22, 23-289, 290-526, 527-640, 641-896, or 897-1239 of SEQ ID NO:13, 1-255, 256-442, 443-626, 627-954, or 955-1297 of SEQ ID NO:14, 1-255, 256-476, 477-680, 681-983, or 984-1326 of SEQ ID NO:15, 1-13, 14-273, 274-346, 347-534, or 535-894 of SEQ ID NO:18, 1-21, 22-277, 278-427, 428-637, 638-751, or 752-898 of SEQ ID NO:22, 1-66, 67-215, 216-425, 426-539, 540-786, or 787-887 of SEQ ID NO:23, 1-25, 26-273, 274-422, 423-632, or 633-748 of SEQ ID NO:24, 1-51, 52-224, 225-393, 394-658, or 659-681 of SEQ ID NO:29, 1-25, 26-281, 282-430, 431-640, 641-754, 755-901, or 902-1001 of SEQ ID NO:31, 1-10, 11-321, or 322-373 of SEQ ID NO:97, 1-57, 58-369, or 370-418 of SEQ ID NO:99, 1-52, 53-173, 174-307, 308-572, or 573-591 of SEQ ID NO:103, 1-24, 25-289, 290-397, 398-628, 629-668, 669-872, or 873-1227 of SEQ ID NO:105, or 1-33, 34-294, 295-337, 338-472, 473-724, or 725-968 of SEQ ID NO:107; (c) the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107 from amino acid residues 1-21, 22-274, or 275-416 of

SEQ ID NO:5, 1-31, 32-308, 309-489, or 490-516 of SEQ ID
NO:6, 1-178 or 179-414 of SEQ ID NO:7, 23-289, 290-526, 527-
640, 641-896, or 897-1239 of SEQ ID NO:13, 1-255, 256-442,
443-626, 627-954, or 955-1297 of SEQ ID NO:14, 1-255, 256-
476, 477-680, 681-983, or 984-1326 of SEQ ID NO:15, 1-13,
14-273, 274-346, 347-534, or 535-894 of SEQ ID NO:18, 1-21,
22-277, 278-427, 428-637, 638-751, or 752-898 of SEQ ID
NO:22, 1-66, 67-215, 216-425, 426-539, 540-786, or 787-887
of SEQ ID NO:23, 1-25, 26-273, 274-422, 423-632, or 633-748
of SEQ ID NO:24, 1-51, 52-224, 225-393, 394-658, or 659-681
of SEQ ID NO:29, 1-25, 26-273, 274-422, 423-632, 633-746,
747-993, or 994-1093 of SEQ ID NO:31, 1-10, 11-321, or 322-
373 of SEQ ID NO:97, 1-57, 58-369, or 370-418 of SEQ ID
NO:99, 1-52, 53-173, 174-307, 308-572, or 573-591 of SEQ ID
NO:103, 1-24, 25-289, 290-397, 398-628, 629-668, 669-872, or
873-1227 of SEQ ID NO:105, or 1-33, 34-294, 295-337, 338-
472, 473-724, or 725-968 of SEQ ID NO:107; or (d) the amino
acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID
NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID
NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID
NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID
NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107,
except that it lacks one or more, but not all, of the
domains selected from the group consisting of a C-terminal
domain, a catalytic domain, an N-terminal domain, a spacer
region, a proline-rich region, a coiled-coil structure
region, an insert, and a C-terminal tail.

The polypeptide can be isolated from a natural source
by methods well-known in the art. The natural source may be
mammalian, preferably human, blood, semen, or tissue, and
the polypeptide may be synthesized using an automated
polypeptide synthesizer. The isolated, enriched, or
purified kinase polypeptide is preferably: a STLK2, STLK3,

STLK4, STLK5, STLK6, or STLK7 polypeptide; a ZC1, ZC2, ZC3, or ZC4 polypeptide; a KHS2 polypeptide; a SULU1 or SULU3 polypeptide; a GEK2 polypeptide; or a PAK4 or PAK5 polypeptide.

5 In some embodiments the invention includes a recombinant kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5. By "recombinant kinase polypeptide" is meant a polypeptide
10 produced by recombinant DNA techniques such that it is distinct from a naturally occurring polypeptide either in its location (e.g., present in a different cell or tissue than found in nature), purity or structure. Generally, such a recombinant polypeptide will be present in a cell in an
15 amount different from that normally observed in nature.

In a fifth aspect, the invention features an antibody (e.g., a monoclonal or polyclonal antibody) having specific binding affinity to a kinase polypeptide or a kinase polypeptide domain or fragment where the polypeptide is selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5. By "specific binding affinity" is meant that the antibody binds to the target kinase polypeptide with greater affinity than it binds to other polypeptides under specified conditions. Antibodies or antibody fragments are polypeptides that contain regions that can bind other polypeptides. The term "specific binding affinity" describes an antibody that binds to a kinase polypeptide with greater affinity than it binds to other polypeptides under specified conditions.
20
25
30

The term "polyclonal" refers to antibodies that are heterogenous populations of antibody molecules derived from the sera of animals immunized with an antigen or an

5

antigenic functional derivative thereof. For the production of polyclonal antibodies, various host animals may be immunized by injection with the antigen. Various adjuvants may be used to increase the immunological response, depending on the host species.

10

"Monoclonal antibodies" are substantially homogenous populations of antibodies to a particular antigen. They may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. Monoclonal antibodies may be obtained by methods known to those skilled in the art (Kohler et al., Nature 256:495-497, 1975, and U.S. Patent No. 4,376,110, both of which are hereby incorporated by reference herein in their entirety including any figures, tables, or drawings).

15

20

The term "antibody fragment" refers to a portion of an antibody, often the hyper variable region and portions of the surrounding heavy and light chains, that displays specific binding affinity for a particular molecule. A hyper variable region is a portion of an antibody that physically binds to the polypeptide target.

25

30

Antibodies or antibody fragments having specific binding affinity to a kinase polypeptide of the invention may be used in methods for detecting the presence and/or amount of kinase polypeptide in a sample by probing the sample with the antibody under conditions suitable for kinase-antibody immunocomplex formation and detecting the presence and/or amount of the antibody conjugated to the kinase polypeptide. Diagnostic kits for performing such methods may be constructed to include antibodies or antibody fragments specific for the kinase as well as a conjugate of a binding partner of the antibodies or the antibodies themselves.

An antibody or antibody fragment with specific binding affinity to a kinase polypeptide of the invention can be isolated, enriched, or purified from a prokaryotic or eukaryotic organism. Routine methods known to those skilled in the art enable production of antibodies or antibody fragments, in both prokaryotic and eukaryotic organisms. Purification, enrichment, and isolation of antibodies, which are polypeptide molecules, are described above.

Antibodies having specific binding affinity to a kinase polypeptide of the invention may be used in methods for detecting the presence and/or amount of kinase polypeptide in a sample by contacting the sample with the antibody under conditions such that an immunocomplex forms and detecting the presence and/or amount of the antibody conjugated to the kinase polypeptide. Diagnostic kits for performing such methods may be constructed to include a first container containing the antibody and a second container having a conjugate of a binding partner of the antibody and a label, such as, for example, a radioisotope. The diagnostic kit may also include notification of an FDA approved use and instructions therefor.

In a sixth aspect, the invention features a hybridoma which produces an antibody having specific binding affinity to a kinase polypeptide or a kinase polypeptide domain, where the polypeptide is selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5. By "hybridoma" is meant an immortalized cell line that is capable of secreting an antibody, for example an antibody to a kinase of the invention. In preferred embodiments, the antibody to the kinase comprises a sequence of amino acids that is able to specifically bind a kinase polypeptide of the invention.

In a seventh aspect, the invention features a kinase polypeptide binding agent able to bind to a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK6, STLK7, STLK5, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5. The binding agent is preferably a purified antibody that recognizes an epitope present on a kinase polypeptide of the invention. Other binding agents include molecules that bind to kinase polypeptides and analogous molecules that bind to a kinase polypeptide. Such binding agents may be identified by using assays that measure kinase binding partner activity, such as those that measure PDGFR activity.

The invention also features a method for screening for human cells containing a kinase polypeptide of the invention or an equivalent sequence. The method involves identifying the novel polypeptide in human cells using techniques that are routine and standard in the art, such as those described herein for identifying the kinases of the invention (e.g., cloning, Southern or Northern blot analysis, in situ hybridization, PCR amplification, etc.).

In an eighth aspect, the invention features methods for identifying a substance that modulates kinase activity comprising the steps of: (a) contacting a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5 with a test substance; (b) measuring the activity of said polypeptide; and (c) determining whether said substance modulates the activity of said polypeptide.

The term "modulates" refers to the ability of a compound to alter the function of a kinase of the invention. A modulator preferably activates or inhibits the activity

of a kinase of the invention depending on the concentration of the compound exposed to the kinase.

The term "activates" refers to increasing the cellular activity of the kinase. The term inhibit refers to 5 decreasing the cellular activity of the kinase. Kinase activity is preferably the interaction with a natural binding partner.

The term "modulates" also refers to altering the function of kinases of the invention by increasing or 10 decreasing the probability that a complex forms between the kinase and a natural binding partner. A modulator preferably increases the probability that such a complex forms between the kinase and the natural binding partner, more preferably increases or decreases the probability that 15 a complex forms between the kinase and the natural binding partner depending on the concentration of the compound exposed to the kinase, and most preferably decreases the probability that a complex forms between the kinase and the natural binding partner.

20 The term "complex" refers to an assembly of at least two molecules bound to one another. Signal transduction complexes often contain at least two protein molecules bound to one another. For instance, a protein tyrosine receptor protein kinase, GRB2, SOS, RAF, and RAS assemble to form a 25 signal transduction complex in response to a mitogenic ligand.

The term "natural binding partner" refers to polypeptides, lipids, small molecules, or nucleic acids that 30 bind to kinases in cells. A change in the interaction between a kinase and a natural binding partner can manifest itself as an increased or decreased probability that the interaction forms, or an increased or decreased concentration of kinase/natural binding partner complex.

The term "contacting" as used herein refers to mixing a solution comprising the test compound with a liquid medium bathing the cells of the methods. The solution comprising the compound may also comprise another component, such as dimethyl sulfoxide (DMSO), which facilitates the uptake of the test compound or compounds into the cells of the methods. The solution comprising the test compound may be added to the medium bathing the cells by utilizing a delivery apparatus, such as a pipet-based device or syringe-based device.

In a ninth aspect, the invention features methods for identifying a substance that modulates kinase activity in a cell comprising the steps of: (a) expressing a kinase polypeptide in a cell, wherein said polypeptide is selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5; (b) adding a test substance to said cell; and (c) monitoring a change in cell phenotype or the interaction between said polypeptide and a natural binding partner.

The term "expressing" as used herein refers to the production of kinases of the invention from a nucleic acid vector containing kinase genes within a cell. The nucleic acid vector is transfected into cells using well known techniques in the art as described herein.

In a tenth aspect, the invention provides methods for treating a disease by administering to a patient in need of such treatment a substance that modulates the activity of a kinase selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5. Preferably, the disease is selected from the group consisting of immune-related diseases and disorders, organ transplantation, myocardial

infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders, and cancer. Most preferably, the immune-related diseases and disorders include, but are not limited to, rheumatoid arthritis, atherosclerosis, and autoimmune disorders.

In preferred embodiments, the invention provides methods for treating or preventing a disease or disorder by administering to a patient in need of such treatment a substance that modulates the activity of a kinase polypeptide selected from the group consisting of ZC1, ZC2, ZC3, ZC4, KHS2, PAK4, and PAK5. Preferably, the disease or disorder is selected from the group consisting of rheumatoid arthritis, atherosclerosis, autoimmune disorders, and organ transplantation. The invention also features methods of treating or preventing a disease or disorder by administering to a patient in need of such treatment a substance that modulates the activity of a kinase polypeptide selected from the group consisting of STLK1, STLK2, STLK3, STLK4, STLK5, STLK6, and STLK7. Preferably the disease or disorder is selected from the group consisting of immune-related diseases and disorders, myocardial infarction, cardiomyopathies, stroke, renal failure, and oxidative stress-related neurodegenerative disorders. Most preferably, the immune-related diseases and disorders are selected from the group consisting of rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplantation.

The invention also features methods of treating or preventing a disease or disorder by administering to a patient in need of such treatment a substance that modulates the activity of a kinase polypeptide selected from the group

consisting of ZC1, ZC2, ZC3, and ZC4. Preferably the disease is selected from the group consisting of immune-related diseases and disorders, cardiovascular disease, and cancer. Most preferably, the immune-related diseases and disorders are selected from the group consisting of rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplantation.

Substances useful for treatment of kinase-related disorders or diseases preferably show positive results in one or more *in vitro* assays for an activity corresponding to treatment of the disease or disorder in question (Examples of such assays are provided in the references in section VI, below; and in Example 7, herein). Examples of substances that can be screened for favorable activity are provided and referenced in section VI, below. The substances that modulate the activity of the kinases preferably include, but are not limited to, antisense oligonucleotides and inhibitors of protein kinases, as determined by methods and screens referenced in section VI and Example 7, below.

The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

The term "therapeutic effect" refers to the inhibition or activation factors causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of

the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (i.e., slowing or stopping) of cell death; (c) inhibition of degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population of cells.

5 Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, or cell survival.

10 Abnormal cell proliferative conditions include cancers such as fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus, and inflammation.

15 Abnormal differentiation conditions include, but are not limited to neurodegenerative disorders, slow wound healing rates, and slow tissue grafting healing rates.

20 Abnormal cell survival conditions relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the 25 function of any one of the protein kinases could lead to cell immortality or premature cell death.

30 The term "aberration", in conjunction with the function of a kinase in a signal transduction process, refers to a kinase that is over- or under-expressed in an organism, mutated such that its catalytic activity is lower or higher than wild-type protein kinase activity, mutated such that it can no longer interact with a natural binding partner, is no longer modified by another protein kinase or protein

phosphatase, or no longer interacts with a natural binding partner.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques, and carrier techniques.

The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism.

The effect of administering a compound on organism function can then be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig, or goat, more preferably a monkey or ape, and most preferably a human.

In an eleventh aspect, the invention features methods for detection of a kinase polypeptide in a sample as a diagnostic tool for diseases or disorders, wherein the method comprises the steps of: (a) contacting the sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5, said probe comprising the nucleic acid sequence encoding the

polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease.

5 In preferred embodiments of the invention, the disease or disorder is selected from the group consisting of rheumatoid arthritis, atherosclerosis, autoimmune disorders, organ transplantation, myocardial infarction, cardiomyopathies, stroke, renal failure, oxidative stress-related neurodegenerative disorders, and cancer. In other preferred embodiments, the kinase polypeptide is selected from the group consisting of PAK4 and PAK5, or the polypeptide is selected from the group consisting of ZC1, ZC2, ZC3, and ZC4, and the disease is cancer.

10 15 The kinase "target region" is the nucleotide base sequence set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:27, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, or SEQ ID NO:106, or the corresponding full-length sequences, a functional derivative thereof, or a fragment thereof to which the nucleic acid probe will specifically hybridize. Specific hybridization indicates that in the presence of other nucleic acids the probe only hybridizes 20 25 detectably with the kinase of the invention's target region.

Putative target regions can be identified by methods well known in the art consisting of alignment and comparison of the most closely related sequences in the database.

30 In preferred embodiments the nucleic acid probe hybridizes to a kinase target region encoding at least 6, 12, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of the sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15,

SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, or the corresponding full-length amino acid sequence, or a functional derivative thereof. Hybridization conditions should be such that hybridization occurs only with the kinase genes in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize.

5 Preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined supra.

10 The diseases for which detection of kinase genes in a sample could be diagnostic include diseases in which kinase nucleic acid (DNA and/or RNA) is amplified in comparison to normal cells. By "amplification" is meant increased numbers of kinase DNA or RNA in a cell compared with normal cells. In normal cells, kinases are typically found as single copy genes. In selected diseases, the chromosomal location of 15 the kinase genes may be amplified, resulting in multiple copies of the gene, or amplification. Gene amplification can lead to amplification of kinase RNA, or kinase RNA can be amplified in the absence of kinase DNA amplification.

20 "Amplification" as it refers to RNA can be the detectable presence of kinase RNA in cells, since in some normal cells there is no basal expression of kinase RNA. In other normal cells, a basal level of expression of kinase exists, therefore in these cases amplification is the detection of at least 1-2-fold, and preferably more, kinase 25 RNA, compared to the basal level.

30 The diseases that could be diagnosed by detection of kinase nucleic acid in a sample preferably include cancers. The test samples suitable for nucleic acid probing methods

of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

In a final aspect, the invention features a method for detection of a kinase polypeptide in a sample as a diagnostic tool for a disease or disorder, wherein the method comprises: (a) comparing a nucleic acid target region encoding the kinase polypeptide in a sample, where the kinase polypeptide is selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5, or one or more fragments thereof, with a control nucleic acid target region encoding the kinase polypeptide, or one or more fragments thereof; and (b) detecting differences in sequence or amount between the target region and the control target region, as an indication of the disease or disorder. Preferably, the disease or disorder is selected from the group consisting of immune-related diseases and disorders, organ transplantation, myocardial infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders, and cancer. Immune-related diseases and disorders include, but are not limited to, those discussed previously.

The term "comparing" as used herein refers to identifying discrepancies between the nucleic acid target region isolated from a sample, and the control nucleic acid target region. The discrepancies can be in the nucleotide sequences, e.g. insertions, deletions, or point mutations,

5

or in the amount of a given nucleotide sequence. Methods to determine these discrepancies in sequences are well-known to one of ordinary skill in the art. The "control" nucleic acid target region refers to the sequence or amount of the sequence found in normal cells, e.g. cells that are not diseased as discussed previously.

10

The invention has been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein. For example, in some instances the nucleotide sequence of the ZC4 kinase polypeptide may not be part of a preferred embodiment.

15

The summary of the invention described above is not limiting and other features and advantages of the invention will be apparent from the following detailed description of the invention, and from the claims.

BRIEF DESCRIPTION OF THE FIGURES

25

Figure 1 shows a multiple sequence alignment of the amino acid sequences of the STE20-STE20 family kinases.

Figure 2 shows a multiple sequence alignment of the amino acid sequences of the STE20-STLK5 family kinases.

Figures 3A and 3B show a multiple sequence alignment of the amino acid sequences of STE20-ZC family kinases.

30

Figure 4 shows a pairwise sequence alignment of STE20-KHS family kinases.

Figure 5 shows a multiple sequence alignment of the amino acid sequences of STE20-SULU family kinases.

Figure 6 shows a pairwise sequence alignment of STE20-GEK family kinases

Figure 7 shows a multiple sequence alignment of the amino acid sequences of STE20-PAK family kinases.

Figures 8A, 8B, 8C, 8D, and 8E show the amino acid sequences of human STLK2, human STLK3, human STLK4, human STLK5, human ZC1, human ZC2, human ZC3, human ZC4, human KHS2, human SULU1, human SULU3, murine SULU3, human GEK2, human PAK4, and human PAK5.

Figures 9A, 9B, 9C, 9D, 9E, 9F, 9G, 9H, 9I, 9J, 9K, 9L, 9M, 9N, 9O, and 9P show the nucleic acid sequences of human STLK2, human STLK3, human STLK4, human STLK5, human ZC1, human ZC2, human ZC3, human ZC4, human KHS2, human SULU1, human SULU3, murine SULU3, human GEK2, human PAK4, and human PAK5.

Figures 10A and 10B show the full-length amino acid sequences of human STLK5 (SEQ ID NO: 97), human PAK5 (SEQ ID NO:103), and human ZC4 (SEQ ID NO:105), as well as the partial amino acid sequences of human full-length STLK6 (SEQ ID NO: 99) and human STLK7 (SEQ ID NO: 101).

Figures 11A, 11B, 11C, and 11D show the full-length nucleic acid sequences of human STLK5 (SEQ ID NO:96), human PAK5 (SEQ ID NO:102), and human ZC4 (SEQ ID NO:104), as well as the partial nucleic acid sequences of human STLK6 (SEQ ID NO: 98) and human STLK7 (SEQ ID NO: 100).

Figure 12 shows a multiple sequence alignment among human SPAK, human STLK6, human STLK7 and full-length human STLK5.

Figure 13 shows a multiple sequence alignment among human PAK1, human PAK4 and human PAK5.

Figures 14A and 14B show a pair-wise sequence alignment between human ZC1 and human ZC4.

Figure 15 shows a pair-wise sequence alignment between LOK1 and full-length GEK2.

DETAILED DESCRIPTION OF THE INVENTION

5 The present invention relates in part to kinase polypeptides, nucleic acids encoding such polypeptides, cells containing such nucleic acids, antibodies to such polypeptides, assays utilizing such polypeptides, and methods relating to all of the foregoing. The present
10 invention is based upon the isolation and characterization of new kinase polypeptides. The polypeptides and nucleic acids may be produced using well-known and standard synthesis techniques when given the sequences presented herein.

15 The recent elucidation of the DNA sequence of *Saccharomyces cerevesiae* has provided the first complete example of the genetic information contained in a simple eukaryotic organism. Analysis of this yeast genome revealed that it contains at least 113 protein kinases. These
20 kinases were further subdivided into several structurally related groups. One of these newly defined groups was termed the STE20-family to represent its founding member STE20, which is a protein kinase involved in the yeast pheromone response pathway that initiates a protein kinase cascade in response to a G-protein mediated signal. *S. cerevesiae* has two additional members of this family, CLA4,
25 and YOL113W (HRA655).

30 Several mammalian homologues have recently been identified that belong to the STE20-family, including SOK-1 (human STE20), GC-kinase, KHS, HPK1, NIK, SLK, GEK, PAK1, PAK65, MST1, and CDC7. Furthermore, the *Drosophila* and the *C. elegans* genome efforts have identified additional protein kinases which belong to the STE20-family, yet have

structurally unique extracatalytic domains, including ZC504.4 and SULU kinases from *C. elegans*, and NINAC of *Drosophila*.

STE20-related protein kinases have been implicated as regulating a variety of cellular responses, including response to growth factors or cytokines, oxidative-, UV-, or irradiation-related stress pathways, inflammatory signals (i.e., TNF α), apoptotic stimuli (i.e., Fas), T and B cell costimulation, the control of cytoskeletal architecture, and cellular transformation. Typically, the STE20-related kinases serve as upstream regulators of MAPK cascades.

Examples include: HPK1, a protein-serine/ threonine kinase (STK) that possesses a STE20-like kinase domain that activates a protein kinase pathway leading to the stress-activated protein kinase SAPK/JNK; PAK1, an STK with an upstream CDC42-binding domain that interacts with Rac and plays a role in cellular transformation through the Ras-MAPK pathway; and murine NIK, which interacts with upstream receptor tyrosine kinases and connects with downstream STE11-family kinases.

The STE20-kinases possess a variety of non-catalytic domains that are believed to interact with upstream regulators. Examples include proline-rich domains for interaction with SH3-containing proteins, or specific domains for interaction with Rac, Rho, and Rab small G-proteins. These interactions may provide a mechanism for cross-talk between distinct biochemical pathways in response to external stimuli such as the activation of a variety of cell surface receptors, including tyrosine kinases, cytokine receptors, TNF receptor, Fas, T cell receptors, CD28, or CD40.

I. The Nucleic Acids of the Invention

Included within the scope of this invention are the functional equivalents of the herein-described isolated nucleic acid molecules. The degeneracy of the genetic code permits substitution of certain codons by other codons that specify the same amino acid and hence would give rise to the same protein. The nucleic acid sequence can vary substantially since, with the exception of methionine and tryptophan, the known amino acids can be coded for by more than one codon. Thus, portions or all of the kinase genes of the invention could be synthesized to give a nucleic acid sequence significantly different from that shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:27, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, and SEQ ID NO:106. The encoded amino acid sequence thereof would, however, be preserved.

In addition, the nucleic acid sequence may comprise a nucleotide sequence which results from the addition, deletion or substitution of at least one nucleotide to the 5'-end and/or the 3'-end of the nucleic acid formula shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:27, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, or SEQ ID NO:106, or a derivative thereof. Any nucleotide or polynucleotide may be used in this regard, provided that its addition, deletion or substitution does not alter the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID

NO:105, or SEQ ID NO:107, which is encoded by the nucleotide sequence. For example, the present invention is intended to include any nucleic acid sequence resulting from the addition of ATG as an initiation codon at the 5'-end of the inventive nucleic acid sequence or its derivative, or from the addition of TTA, TAG or TGA as a termination codon at the 3'-end of the inventive nucleotide sequence or its derivative. Moreover, the nucleic acid molecule of the present invention may, as necessary, have restriction endonuclease recognition sites added to its 5'-end and/or 3'-end.

Such functional alterations of a given nucleic acid sequence afford an opportunity to promote secretion and/or processing of heterologous proteins encoded by foreign nucleic acid sequences fused thereto. All variations of the nucleotide sequence of the kinase genes of the invention and fragments thereof permitted by the genetic code are, therefore, included in this invention.

Further, it is possible to delete codons or to substitute one or more codons with codons other than degenerate codons to produce a structurally modified polypeptide, but one which has substantially the same utility or activity as the polypeptide produced by the unmodified nucleic acid molecule. As recognized in the art, the two polypeptides are functionally equivalent, as are the two nucleic acid molecules that give rise to their production, even though the differences between the nucleic acid molecules are not related to the degeneracy of the genetic code.

30

Mammalian STLK2

The full-length human STLK2 cDNA (SEQ ID NO:1) is 3268 bp long and consists of a 1248 bp open reading frame (ORF)

5

10

flanked by a 181 bp 5' untranslated region (UTR; 1-181) and a 1784 bp 3' UTR (1433-3216) that is followed by a 52 nucleotide polyadenylated region. A polyadenylation signal (AATAAA) is found at positions (3193-3198). The sequence flanking the first ATG conforms to the Kozak consensus (Kozak, M., Nucleic Acids Res. 15, 8125-8148 (1987)) for an initiating methionine, and is believed to be the translational start site for STLK2. Furthermore, human STLK2 and the related SOK-1 and MST3 proteins conserve the amino acid sequence immediately following this presumed initiating methionine.

Several EST fragments span the complete STLK2 sequence with AA191319 at the 5' end and W16504 at the 3' end.

15

Mammalian STLK3

20

The partial human STLK3 cDNA (SEQ ID NO:2) is 3030 bp long and consists of a 1548 bp ORF flanked by a 1476 bp 3' UTR (1550-3025) and a 5 nucleotide polyadenylated region. A potential polyadenylation signal (AATAAA) begins at position 3004. Since the coding region is open throughout the 5' extent of this sequence, this is apparently a partial cDNA clone lacking the N-terminal start methionine.

25

Multiple EST fragments span the complete STLK3 sequence with AA278967 at the 5' end and AA628477 and others at the 3' end.

Mammalian STLK4

30

The partial human STLK4 cDNA (SEQ ID NO:3) is 3857 bp long and consists of a 1242 bp ORF flanked by a 2596 bp 3' UTR (1244-3839) and an 18 nucleotide polyadenylated region. A potential polyadenylation signal (AATAAA) is found at positions 2181-3822. Since the coding region is open throughout the 5' extent of this sequence, this is

apparently a partial cDNA clone lacking the N-terminal start methionine. A near full-length murine STLK4 cDNA is represented in the 1773 bp EST AA117438. It extends an additional 21 nucleotides 5' of the human STLK4 consensus, but since its coding region is open throughout the 5' extent of the sequence, this is also apparently a partial cDNA clone lacking the N-terminal start methionine.

5 Several EST fragments span the complete STLK3 sequence with AA297759 at the 5' end and AA100484 and others at the
10 3' end.

Mammalian STLK5

15 The full-length human STLK5 cDNA (SEQ ID NO:96) is 2110 bp long and consists of a 1119 bp ORF flanked by a 229 bp 5' UTR and a 762 bp 3' UTR. The sequence flanking the first ATG conforms to the Kozak consensus (supra) for an initiating methionine, and is believed to be the translational start site for STLK5. Several EST fragments span the complete
20 STLK5 sequence with AA297059 and F07734 at the 5' end, and R46686 and F03423 and others at the 3' end.

Mammalian STLK6

25 The full-length human STLK6 cDNA (SEQ ID NO:98) is 2,001 bp long and consists of a 1,254 bp ORF flanked by a 75 bp 5' UTR and a 673 bp 3' UTR. The sequence flanking the first ATG conforms to the Kozak consensus (supra) for an initiating methionine, and is believed to be the translational start site for STLK6.

30 Mammalian STLK7

The partial human STLK7 cDNA (SEQ ID NO:100) is 311 bp long and consists of a 309 bp ORF. Since the coding region is open throughout both the 5' and 3' extent of this

sequence, this is apparently a partial cDNA clone lacking the N-terminal start methionine and C-terminal stop codon.

Mammalian ZC1

5 The full-length human ZC1 cDNA (SEQ ID NO:9) is 3798 bp long and consists of a 3717 bp ORF (7-3723) flanked by a 6 bp 5' UTR and a 75 bp (3724-3798) 3' UTR. No polyadenylation signal (AATAAA) or polyadenylated region are present in the 3'UTR. The sequence flanking the first ATG
10 conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for human ZC1.

15 Multiple EST fragments (W81656) match the 3' end of the human ZC1 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

Mammalian ZC2

20 The partial human ZC2 cDNA (SEQ ID NO:10) is 4055 bp long and consists of a 3891 bp ORF (1-3891) and a 164 bp (3892-4055) 3' UTR. Since the coding region is open throughout the 5' extent of this sequence, this is apparently a partial cDNA clone lacking the N-terminal start methionine. No polyadenylation signal (AATAAA) or
25 polyadenylated region are present in the 3'UTR.

Multiple EST fragments (R51245) match the 3' end of the human ZC2 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

30

Mammalian ZC3

The partial human ZC3 cDNA (SEQ ID NO:11) is 4133 bp long and consists of a 3978 bp ORF (1-3978) and a 152 bp

(3979-4133) 3'UTR region. Since the coding region is open throughout the 5' extent of this sequence, this is apparently a partial cDNA clone lacking the N-terminal start methionine. No polyadenylation signal (AATAAA) or polyadenylated region are present in the 3'UTR.

Multiple EST fragments (R54563) match the 3'end of the human ZC3 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5'.end.

10

Mammalian ZC4

The full-length human ZC4 cDNA (SEQ ID NO:104) is 3,684 bp long and was originally assembled from X chromosome genomic DNA sequence.

15

Multiple EST fragments (R98571) match the 3'end of the human ZC4 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end. ZC4 gene is also contained within the human genomic clone Z83850.

20

Mammalian KHS2

The full-length human KHS2 cDNA (SEQ ID NO:17) is 4023 bp long and consists of a 2682 bp ORF (6-2687) flanked by a 5 bp (1-5) 5'UTR and a 1336 bp (2688-4023) 3' UTR. A potential polyadenylation signal (AATAAA) is found at positions 4008-4013. No polyadenylated region is present in the 3'UTR. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for human KHS2.

30

Multiple EST fragments match the 5'end (AA446022) as well as the 3' end (R37625) of the human KHS2 gene.

Mammalian SULU1

The full-length human SULU1 cDNA (SEQ ID NO:19) is 4177 bp long and consists of a 2694 bp ORF (415-3108) flanked by a 414 bp (1-414) 5'UTR and a 1069 bp (3109-4177) 3' UTR
5 followed by a 19 nucleotide polyadenylated region. A potential polyadenylation signal (AATAAA) is found at positions 4164-4169. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start
10 site for human SULU1.

Multiple EST fragments match the 5'end (N27153) as well as the 3' end (R90908) of the human SULU1 gene.

Mammalian (Murine) SULU3

15 The partial murine SULU3 cDNA (SEQ ID NO:21) is 2249 bp long and consists of a 2244 bp ORF (6-2249) flanked by a 5 bp (1-5) 5'UTR. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start
20 site for murine SULU3. The 3' end of the murine SULU3 cDNA shares 90% DNA sequence identity over 1620 nucleotides with human SULU3, suggesting that these two genes are functional orthologues.

One EST fragment (AA446022) matches the 3' end of the
25 partial murine SULU3 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

Mammalian (Human) SULU3

30 The partial human SULU3 cDNA (SEQ ID NO:20) is 3824 bp long and consists of a 2358 bp ORF (2-2359) flanked by a 1465 bp (2360-3824) 3'UTR followed by a 19 nucleotide polyadenylated region. A potential polyadenylation signal

(AATAAA) is found at positions 2602-2607. Since the coding region is open throughout the 5' extent of this sequence, this is apparently a partial cDNA clone lacking the N-terminal start methionine. The 5' end of the human SULU3 cDNA shares 90% DNA sequence identity over 1620 nucleotides with murine SULU3, suggesting that these two genes are functional orthologues.

Multiple EST fragments (R02283) match the 3'end of the human SULU3 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

Mammalian GEK2

The full-length human GEK2 cDNA (SEQ ID NO:106) is 2962 bp long and consists of a 2737 bp ORF (59-2795) flanked by a 58 bp (1-58) 5'UTR. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for human GEK2.

Multiple EST fragments (AA465671) match the 5'end, but at the time of filing, the inventors believe that only one (AA380492) matches the 3' end of the human GEK2 gene.

Mammalian PAK4

The full-length human PAK4 cDNA (SEQ ID NO:27) is 3604 bp long and consists of a 2043 bp ORF (143-2185) flanked by a 142 bp (1-142) 5'UTR and a 1419 3' UTR followed by a 22 nucleotide polyadenylated region. A potential polyadenylation signal (AATTA_n) is found at positions 3582-3588. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for human PAK4.

Multiple EST fragments (AA535791) match the 3' end of the human PAK4 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

5

Mammalian PAK5

The full-length human PAK5 cDNA (SEQ ID NO:102) is 2806 bp long and consists of a 1773 bp ORF flanked by a 201 bp 5' UTR and a 833 bp 3' UTR. The sequence flanking the first ATG conforms to the Kozak consensus (supra) for an initiating methionine, and is believed to be the translational start site for PAK5.

10

Multiple EST fragments (AA442867) match the 3' end of the human PAK5 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

15

II. Nucleic Acid Probes, Methods, and Kits for Detection of STE20-Related Kinases.

20

A nucleic acid probe of the present invention may be used to probe an appropriate chromosomal or cDNA library by usual hybridization methods to obtain other nucleic acid molecules of the present invention. A chromosomal DNA or cDNA library may be prepared from appropriate cells according to recognized methods in the art (cf. "Molecular Cloning: A Laboratory Manual", second edition, Cold Spring Harbor Laboratory, Sambrook, Fritsch, & Maniatis, eds., 1989).

25

In the alternative, chemical synthesis can be carried out in order to obtain nucleic acid probes having nucleotide sequences which correspond to N-terminal and C-terminal portions of the amino acid sequence of the polypeptide of interest. The synthesized nucleic acid probes may be used

30

as primers in a polymerase chain reaction (PCR) carried out in accordance with recognized PCR techniques, essentially according to PCR Protocols, "A Guide to Methods and Applications", Academic Press, Michael, et al., eds., 1990, 5 utilizing the appropriate chromosomal or cDNA library to obtain the fragment of the present invention.

One skilled in the art can readily design such probes based on the sequence disclosed herein using methods of computer alignment and sequence analysis known in the art 10 ("Molecular Cloning: A Laboratory Manual", 1989, supra). The hybridization probes of the present invention can be labeled by standard labeling techniques such as with a radiolabel, enzyme label, fluorescent label, biotin-avidin label, chemiluminescence, and the like. After 15 hybridization, the probes may be visualized using known methods.

The nucleic acid probes of the present invention include RNA, as well as DNA probes, such probes being generated using techniques known in the art. The nucleic 20 acid probe may be immobilized on a solid support. Examples of such solid supports include, but are not limited to, plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, and acrylic resins, such as polyacrylamide and latex beads. Techniques for coupling 25 nucleic acid probes to such solid supports are well known in the art.

The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. 30 The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are

well known in the art and can be readily adapted in order to obtain a sample which is compatible with the method utilized.

One method of detecting the presence of nucleic acids of the invention in a sample comprises (a) contacting said sample with the above-described nucleic acid probe under conditions such that hybridization occurs, and (b) detecting the presence of said probe bound to said nucleic acid molecule. One skilled in the art would select the nucleic acid probe according to techniques known in the art as described above. Samples to be tested include but should not be limited to RNA samples of human tissue.

A kit for detecting the presence of nucleic acids of the invention in a sample comprises at least one container means having disposed therein the above-described nucleic acid probe. The kit may further comprise other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound nucleic acid probe. Examples of detection reagents include, but are not limited to radiolabelled probes, enzymatic labeled probes (horseradish peroxidase, alkaline phosphatase), and affinity labeled probes (biotin, avidin, or streptavidin).

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the probe or primers used in the

assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, and the like), and containers which contain the reagents used to detect the hybridized probe, bound antibody, amplified product, or the like. One skilled in the art will readily recognize that the nucleic acid probes described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

10 III. DNA Constructs Comprising a STE20-Related Nucleic Acid Molecule and Cells Containing These Constructs.

The present invention also relates to a recombinant DNA molecule comprising, 5' to 3', a promoter effective to initiate transcription in a host cell and the above-described nucleic acid molecules. In addition, the present invention relates to a recombinant DNA molecule comprising a vector and an above-described nucleic acid molecule. The present invention also relates to a nucleic acid molecule comprising a transcriptional region functional in a cell, a sequence complementary to an RNA sequence encoding an amino acid sequence corresponding to the above-described polypeptide, and a transcriptional termination region functional in said cell. The above-described molecules may be isolated and/or purified DNA molecules.

25 The present invention also relates to a cell or organism that contains an above-described nucleic acid molecule and thereby is capable of expressing a polypeptide.

30 The polypeptide may be purified from cells which have been altered to express the polypeptide. A cell is said to be "altered to express a desired polypeptide" when the cell, through genetic manipulation, is made to produce a protein which it normally does not produce or which the cell normally produces at lower levels. One skilled in the art

can readily adapt procedures for introducing and expressing either genomic, cDNA, or synthetic sequences into either eukaryotic or prokaryotic cells.

A nucleic acid molecule, such as DNA, is said to be "capable of expressing" a polypeptide if it contains nucleotide sequences which contain transcriptional and translational regulatory information and such sequences are "operably linked" to nucleotide sequences which encode the polypeptide. An operable linkage is a linkage in which the regulatory DNA sequences and the DNA sequence sought to be expressed are connected in such a way as to permit gene sequence expression. The precise nature of the regulatory regions needed for gene sequence expression may vary from organism to organism, but shall in general include a promoter region which, in prokaryotes, contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when transcribed into RNA, will signal synthesis initiation. Such regions will normally include those 5'-non-coding sequences involved with initiation of transcription and translation, such as the TATA box, capping sequence, CAAT sequence, and the like.

If desired, the non-coding region 3' to the sequence encoding a kinase of the invention may be obtained by the above-described methods. This region may be retained for its transcriptional termination regulatory sequences, such as termination and polyadenylation. Thus, by retaining the 3'-region naturally contiguous to the DNA sequence encoding a kinase of the invention, the transcriptional termination signals may be provided. Where the transcriptional termination signals are not satisfactorily functional in the expression host cell, then a 3' region functional in the host cell may be substituted.

Two DNA sequences (such as a promoter region sequence and a sequence encoding a kinase of the invention) are said to be operably linked if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region sequence to direct the transcription of a gene sequence encoding a kinase of the invention, or (3) interfere with the ability of the gene sequence of a kinase of the invention to be transcribed by the promoter region sequence. Thus, a promoter region would be operably linked to a DNA sequence if the promoter were capable of effecting transcription of that DNA sequence. Thus, to express a gene encoding a kinase of the invention, transcriptional and translational signals recognized by an appropriate host are necessary.

The present invention encompasses the expression of a gene encoding a kinase of the invention (or a functional derivative thereof) in either prokaryotic or eukaryotic cells. Prokaryotic hosts are, generally, very efficient and convenient for the production of recombinant proteins and are, therefore, one type of preferred expression system for kinases of the invention. Prokaryotes most frequently are represented by various strains of *E. coli*. However, other microbial strains may also be used, including other bacterial strains.

In prokaryotic systems, plasmid vectors that contain replication sites and control sequences derived from a species compatible with the host may be used. Examples of suitable plasmid vectors may include pBR322, pUC118, pUC119 and the like; suitable phage or bacteriophage vectors may include γ gt10, γ gt11 and the like; and suitable virus vectors may include pMAM-neo, pKRC and the like. Preferably, the

selected vector of the present invention has the capacity to replicate in the selected host cell.

5 Recognized prokaryotic hosts include bacteria such as *E. coli*, *Bacillus*, *Streptomyces*, *Pseudomonas*, *Salmonella*, *Serratia*, and the like. However, under such conditions, the polypeptide will not be glycosylated. The prokaryotic host must be compatible with the replicon and control sequences in the expression plasmid.

To express a kinase of the invention (or a functional derivative thereof) in a prokaryotic cell, it is necessary 10 to operably link the sequence encoding the kinase of the invention to a functional prokaryotic promoter. Such promoters may be either constitutive or, more preferably, regulatable (i.e., inducible or derepressible). Examples of 15 constitutive promoters include the *int* promoter of bacteriophage λ , the *bla* promoter of the β -lactamase gene sequence of pBR322, and the *cat* promoter of the chloramphenicol acetyl transferase gene sequence of pPR325, and the like. Examples of inducible prokaryotic promoters 20 include the major right and left promoters of bacteriophage λ (P_L and P_R), the *trp*, *recA*, *lacZ*, *lacI*, and *gal* promoters of *E. coli*, the α -amylase (Ulmanen et al., J. Bacteriol. 162:176-182, 1985) and the ζ -28-specific promoters of *B. subtilis* (Gilman et al., Gene Sequence 32:11-20, 1984), the 25 promoters of the bacteriophages of *Bacillus* (Gryczan, In: The Molecular Biology of the Bacilli, Academic Press, Inc., NY, 1982), and *Streptomyces* promoters (Ward et al., Mol. Gen. Genet. 203:468-478, 1986). Prokaryotic promoters are reviewed by Glick (Ind. Microbiol. 1:277-282, 1987), 30 Cenatiempo (Biochimie 68:505-516, 1986), and Gottesman (Ann. Rev. Genet. 18:415-442, 1984).

Proper expression in a prokaryotic cell also requires the presence of a ribosome-binding site upstream of the gene

sequence-encoding sequence. Such ribosome-binding sites are disclosed, for example, by Gold et al. (Ann. Rev. Microbiol. 35:365-404, 1981). The selection of control sequences, expression vectors, transformation methods, and the like, are dependent on the type of host cell used to express the gene. As used herein, "cell", "cell line", and "cell culture" may be used interchangeably and all such designations include progeny. Thus, the words "transformants" or "transformed cells" include the primary subject cell and cultures derived therefrom, without regard to the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. However, as defined, mutant progeny have the same functionality as that of the originally transformed cell.

Host cells which may be used in the expression systems of the present invention are not strictly limited, provided that they are suitable for use in the expression of the kinase polypeptide of interest. Suitable hosts may often include eukaryotic cells. Preferred eukaryotic hosts include, for example, yeast, fungi, insect cells, mammalian cells either *in vivo*, or in tissue culture. Mammalian cells which may be useful as hosts include HeLa cells, cells of fibroblast origin such as VERO or CHO-K1, or cells of lymphoid origin and their derivatives. Preferred mammalian host cells include SP2/0 and J558L, as well as neuroblastoma cell lines such as IMR 332, which may provide better capacities for correct post-translational processing.

In addition, plant cells are also available as hosts, and control sequences compatible with plant cells are available, such as the cauliflower mosaic virus 35S and 19S, and nopaline synthase promoter and polyadenylation signal sequences. Another preferred host is an insect cell, for

example the *Drosophila* larvae. Using insect cells as hosts, the *Drosophila* alcohol dehydrogenase promoter can be used (Rubin, Science 240:1453-1459, 1988). Alternatively, baculovirus vectors can be engineered to express large amounts of kinases of the invention in insect cells (Jasny, Science 238:1653, 1987; Miller et al., In: Genetic Engineering, Vol. 8, Plenum, Setlow et al., eds., pp. 277-297, 1986).

Any of a series of yeast expression systems can be utilized which incorporate promoter and termination elements from the actively expressed sequences coding for glycolytic enzymes that are produced in large quantities when yeast are grown in mediums rich in glucose. Known glycolytic gene sequences can also provide very efficient transcriptional control signals. Yeast provides substantial advantages in that it can also carry out post-translational modifications.

A number of recombinant DNA strategies exist utilizing strong promoter sequences and high copy number plasmids which can be utilized for production of the desired proteins in yeast. Yeast recognizes leader sequences on cloned mammalian genes and secretes peptides bearing leader sequences (i.e., pre-peptides). Several possible vector systems are available for the expression of kinases of the invention in a mammalian host.

A wide variety of transcriptional and translational regulatory sequences may be employed, depending upon the nature of the host. The transcriptional and translational regulatory signals may be derived from viral sources, such as adenovirus, bovine papilloma virus, cytomegalovirus, simian virus, or the like, where the regulatory signals are associated with a particular gene sequence which has a high level of expression. Alternatively, promoters from mammalian expression products, such as actin, collagen,

myosin, and the like, may be employed. Transcriptional initiation regulatory signals may be selected which allow for repression or activation, so that expression of the gene sequences can be modulated. Of interest are regulatory signals which are temperature-sensitive so that by varying the temperature, expression can be repressed or initiated, or are subject to chemical (such as metabolite) regulation.

5 Expression of kinases of the invention in eukaryotic hosts requires the use of eukaryotic regulatory regions.

10 Such regions will, in general, include a promoter region sufficient to direct the initiation of RNA synthesis.

Preferred eukaryotic promoters include, for example, the promoter of the mouse metallothionein I gene sequence (Hamer et al., J. Mol. Appl. Gen. 1:273-288, 1982); the TK promoter of Herpes virus (McKnight, Cell 31:355-365, 1982); the SV40 early promoter (Benoist et al., Nature (London) 290:304-31, 1981); and the yeast gal4 gene sequence promoter (Johnston et al., Proc. Natl. Acad. Sci. (USA) 79:6971-6975, 1982; Silver et al., Proc. Natl. Acad. Sci. (USA) 81:5951-5955, 15 20 20 1984).

Translation of eukaryotic mRNA is initiated at the codon which encodes the first methionine. For this reason, it is preferable to ensure that the linkage between a eukaryotic promoter and a DNA sequence which encodes a kinase of the invention (or a functional derivative thereof) does not contain any intervening codons which are capable of encoding a methionine (i.e., AUG). The presence of such codons results either in the formation of a fusion protein (if the AUG codon is in the same reading frame as the kinase of the invention coding sequence) or a frame-shift mutation (if the AUG codon is not in the same reading frame as the kinase of the invention coding sequence).

A nucleic acid molecule encoding a kinase of the invention and an operably linked promoter may be introduced into a recipient prokaryotic or eukaryotic cell either as a nonreplicating DNA or RNA molecule, which may either be a linear molecule or, more preferably, a closed covalent circular molecule. Since such molecules are incapable of autonomous replication, the expression of the gene may occur through the transient expression of the introduced sequence.

Alternatively, permanent expression may occur through the integration of the introduced DNA sequence into the host chromosome.

A vector may be employed which is capable of integrating the desired gene sequences into the host cell chromosome. Cells which have stably integrated the introduced DNA into their chromosomes can be selected by also introducing one or more markers which allow for selection of host cells which contain the expression vector.

The marker may provide for prototrophy to an auxotrophic host, biocide resistance, e.g., antibiotics, or heavy metals, such as copper, or the like. The selectable marker gene sequence can either be directly linked to the DNA gene sequences to be expressed, or introduced into the same cell by co-transfection. Additional elements may also be needed for optimal synthesis of mRNA. These elements may include splice signals, as well as transcription promoters, enhancers, and termination signals. cDNA expression vectors incorporating such elements include those described by Okayama (Mol. Cell. Biol. 3:280-, 1983).

The introduced nucleic acid molecule can be incorporated into a plasmid or viral vector capable of autonomous replication in the recipient host. Any of a wide variety of vectors may be employed for this purpose. Factors of importance in selecting a particular plasmid or

viral vector include: the ease with which recipient cells that contain the vector may be recognized and selected from those recipient cells which do not contain the vector; the number of copies of the vector which are desired in a particular host; and whether it is desirable to be able to "shuttle" the vector between host cells of different species.

Preferred prokaryotic vectors include plasmids such as those capable of replication in *E. coli* (such as, for example, pBR322, ColEl, pSC101, pACYC 184, π VX; "Molecular Cloning: A Laboratory Manual", 1989, supra). *Bacillus* plasmids include pC194, pC221, pT127, and the like (Gryczan, In: *The Molecular Biology of the Bacilli*, Academic Press, NY, pp. 307-329, 1982). Suitable *Streptomyces* plasmids include pIJ101 (Kendall et al., *J. Bacteriol.* 169:4177-4183, 1987), and *streptomyces* bacteriophages such as ϕ C31 (Chater et al., In: *Sixth International Symposium on Actinomycetales Biology*, Akademiai Kaido, Budapest, Hungary, pp. 45-54, 1986). *Pseudomonas* plasmids are reviewed by John et al. (*Rev. Infect. Dis.* 8:693-704, 1986), and Izaki (Jpn. *J. Bacteriol.* 33:729-742, 1978).

Preferred eukaryotic plasmids include, for example, BPV, vaccinia, SV40, 2-micron circle, and the like, or their derivatives. Such plasmids are well known in the art (Botstein et al., *Miami Wntr. Symp.* 19:265-274, 1982; Broach, In: "The Molecular Biology of the Yeast *Saccharomyces*: Life Cycle and Inheritance", Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, p. 445-470, 1981; Broach, *Cell* 28:203-204, 1982; Bollon et al., *J. Clin. Hematol. Oncol.* 10:39-48, 1980; Maniatis, In: *Cell Biology: A Comprehensive Treatise*, Vol. 3, *Gene Sequence Expression*, Academic Press, NY, pp. 563-608, 1980).

Once the vector or nucleic acid molecule containing the construct(s) has been prepared for expression, the DNA construct(s) may be introduced into an appropriate host cell by any of a variety of suitable means, i.e., transformation, transfection, conjugation, protoplast fusion, electroporation, particle gun technology, calcium phosphate-precipitation, direct microinjection, and the like. After the introduction of the vector, recipient cells are grown in a selective medium, which selects for the growth of vector-containing cells. Expression of the cloned gene(s) results in the production of a kinase of the invention, or fragments thereof. This can take place in the transformed cells as such, or following the induction of these cells to differentiate (for example, by administration of bromodeoxyuracil to neuroblastoma cells or the like). A variety of incubation conditions can be used to form the peptide of the present invention. The most preferred conditions are those which mimic physiological conditions.

IV. The Proteins of the Invention

A variety of methodologies known in the art can be utilized to obtain the polypeptides of the present invention. The polypeptides may be purified from tissues or cells that naturally produce the polypeptides. Alternatively, the above-described isolated nucleic acid fragments could be used to express the kinases of the invention in any organism. The samples of the present invention include cells, protein extracts or membrane extracts of cells, or biological fluids. The samples will vary based on the assay format, the detection method, and the nature of the tissues, cells or extracts used as the sample.

Any eukaryotic organism can be used as a source for the polypeptides of the invention, as long as the source organism naturally contains such polypeptides. As used herein, "source organism" refers to the original organism from which the amino acid sequence of the subunit is derived, regardless of the organism the subunit is expressed in and ultimately isolated from.

One skilled in the art can readily follow known methods for isolating proteins in order to obtain the polypeptides free of natural contaminants. These include, but are not limited to: size-exclusion chromatography, HPLC, ion-exchange chromatography, and immuno-affinity chromatography.

Mammalian STLK2

Analysis of the deduced amino acid sequence predicts STLK2 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. STLK2 contains a 21 amino acid N-terminal domain, a 253 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, followed by a 142 amino acid C-terminal domain.

STLK2 is most closely related to human STE20-subfamily kinases, MST3 (GB:AF024636) and SOK-1 (GB:X99325) and a *C. elegans* kinase yk34b11.5 (GB:U53153) sharing 72.7%, 68.7%, and 69.3% amino acid identity, respectively.

The 21 amino acid N-terminal domain of human STLK2 is 71.4% identical to the N-terminus of MST3 (GB:AF024636). Human STLK2 lacks a glycine residue at position 2, and is therefore unlikely to undergo myristylation. A Smith-Waterman search of the nonredundant protein database does not reveal any significant homologies that might suggest a potential function for this domain.

The 253 amino acid catalytic domain of human STLK2 is most related to human SOK-1 (X99325), MST3 (GB:AF024636), *C. elegans* yk32b11.5 (GB:U53153), and STLK3 (SEQ ID NO:6) sharing 88.9%, 87.4%, 78.3%, and 49% amino identity respectively, placing it in the STLK-subfamily of STE20-related kinases. The STLK2 kinase domain displayed lesser homology to other STE20-related kinases including: 55.9% to human MST2 (GB:U26424), 49.2% to human GCK (GB:U07349), 49.2% to human KHS1 (GB:U77129), and 44.2% to human HPK1 (GB:U66464). The activation loop of human STLK2 catalytic domain is identical to that of human SOK-1 and MST3 including the presence of four potential threonine phosphorylation sites that could serve an autoregulatory role on kinase activity.

The 142 amino acid C-terminal domain of human STLK2 is most related to human SOK-1 (X99325), MST3 (GB:AF024636), and *C. elegans* yk32b11.5 (GB:U53153), sharing 39.9%, 39.9%, and 33.3% amino acid identity, respectively. This C-terminal domain shares some significant amino acid similarity to the C-terminal domains of the related human STLK3 (SEQ ID NO:6) and STLK4 (SEQ ID NO:7).

The C-terminus of the related human SOK-1 (GB:X99325) kinase has been shown to be inhibitory to the catalytic activity of this kinase (Pombo, C.M., Bonventre, J.V., Molnar, A., Kyriakis, J. and Force, T. EMBO J. 15, 4537-4546 (1996)). Based on the sequence identity between the C-termini of human SOK-1 (GB:X99325) and human STLK2 (39.2%), the C-terminus of human STLK2 may also function as an inhibitory domain for its kinase.

30

Mammalian STLK3

The 3030 bp human STLK3 nucleotide sequence of the partial cDNA clone encodes a polypeptide of 516 amino acids

(SEQ ID NO:6) with a predicted molecular mass of 56,784 daltons. Analysis of the deduced amino acid sequence predicts STLK3 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain, however the cDNA clone lacks an initiating ATG, so the full extent of its N-terminus is not known. STLK3 contains a 31 amino acid N-terminal domain, a 277 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, followed by a 181 amino acid C-terminal domain containing a 25 amino acid insert and a 27 amino acid tail relative to the sequence of human STLK2.

10 STLK3 is most closely related to human STE20-subfamily kinases, STLK4 (SEQ ID NO:7), MST3 (GB:AF024636), SOK-1 (GB:X99325) and STLK2 (SEQ ID NO:5) sharing 71.1%, 37.6%, 15 38.1%, and 38.4% amino acid identity respectively.

The 31 amino acid N-terminal domain of human STLK3 lacked any significant amino acid sequence homologies using a Smith-Waterman search of the nonredundant protein database, other than sequence similarity to proline-alanine repeats.

20 The 277 amino acid catalytic domain of human STLK3 is most related to human STLK4 (SEQ ID NO:7), SOK-1 (GB:X99325), MST3 (GB:AF024636), and STLK2 (SEQ ID NO:5) sharing 88.2%, 49.2%, 49%, and 49% amino acid identity, 25 respectively. It also shares strong homology to other STKs from lower organisms including 51.7% to *A. thaliana* (GB: AC002343), 43.1% to *A. thaliana* (GB: Z97336), 42.1% to *A. thaliana* (GB: U96613), and 43.3% to *C. elegans* (GB: U53153). 30 The activation loop of the human STLK3 catalytic domain conserves three potential threonine phosphorylation sites with other members of the STLK-subfamily of STE20-related kinases (human STE20, MST3, STLK2, STLK4) that could serve an autoregulatory role on kinase activity.

The 181 amino acid C-terminal domain of human STLK3 shares 55.5% amino acid identity to human STLK4 (SEQ ID NO:7), and is 100% identical to a partial human cDNA DCHT (GB:AF017635). The C-terminal domain of human STLK3 contains a 26 amino acid insert relative to human STE20. A similar (87.5% amino acid identity) 26 amino acid insert is also present in human STLK4.

The 27 amino acid C-terminal tail of human STLK3 shares 77.8% amino acid identity to human STLK4, but is absent from other STLK-family members. This high degree of homology between the C-tail of two STLK-family members suggests they may be involved in an as yet unidentified protein-protein interaction.

The weak sequence homology between the C-termini of human STLK3 and STE20, suggests it may also function as an inhibitory domain for its kinase.

Mammalian STLK4

The 3857 bp human STLK4 nucleotide sequence of the partial cDNA clone encodes a polypeptide of 414 amino acids (SEQ ID NO:7) with a predicted molecular mass of 45,451 daltons. Analysis of the deduced amino acid sequence predicts STLK4 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain, however the cDNA clone lacks an initiating ATG, so the full extent of its N-terminus is not known. The partial STLK4 protein sequence contains a 178 amino acid catalytic domain corresponding to the C-terminal motifs VI-XI of a serine/threonine kinase, followed by a 236 amino acid C-terminal domain containing two inserts of 25 and 41 amino acids each, relative to the sequence of human STLK2.

STLK4 is most closely related to human STE20-subfamily kinases, STLK3 (SEQ ID NO 6), MST3 (GB:AF024636), STLK2

(SEQ ID NO:5), and SOK-1 (GB:X99325) sharing 71.0%, 46.8%, 43.9%, and 37.7% amino acid identity, respectively.

The 178 amino acid catalytic domain of human STLK4 is most related to human STLK3 (SEQ ID NO. 7), SOK-1 (GB:X99325), MST3 (GB:AF024636), STLK2 (SEQ ID NO:5), and MST1 (GB:U18297), sharing 88.2%, 54.2%, 54.0%, 53.7 and 45.7% amino acid identity, respectively. It also shares strong homology to other STKs from lower organisms including 56.9% to *A. thaliana* (GB: AC002343), 52.5% to *C. elegans* (GB: U53153), 46.2% to *A. thaliana* (GB: Z97336) and 45.7% to *A. thaliana* (GB: U96613). The activation loop of the human STLK4 catalytic domain conserves three potential threonine phosphorylation sites with other members of the STLK-subfamily of STE20-related kinases (human STE20, MST3, STLK2 and STLK3) that could serve an autoregulatory role on kinase activity.

The 236 amino acid C-terminal domain of human STLK4 shares 58.1% amino acid identity to both human STLK3 (SEQ ID NO:6) and to a partial human cDNA, DCHT (GB:AF017635). The C-terminal domain of human STLK4 contains a 25 amino acid insert relative to human SOK-1 and shares 87.5% amino acid identity to an insert present in human STLK3.

The weak sequence homology between the C-termini of human STLK4 and STE20, suggests it may also function as an inhibitory domain for its kinase.

Mammalian STLK5

The full-length 2110 bp human STLK5 cDNA encodes a polypeptide of 373 amino acids (SEQ ID NO:57) with a predicted molecular mass of 41,700 daltons. Analysis of the deduced amino acid sequence predicts STLK5 to be an intracellular STE20-subfamily kinase, lacking both a signal sequence and transmembrane domain. STLK5 contains a 10

amino acid N-terminal domain, a 311 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, and a 52 amino acid C-terminal domain.

5 STLK5 is most closely related to the human STE20-subfamily kinases STLK6 (SEQ ID No. 99) and SPAK (AFO99989), sharing 51% and 33% amino acid identity, respectively, over its full extent. It also shares significant homology to database entries from *Arabidopsis thaliana* (GB:AC002343) and
10 C.elegans (GB:AL023843, GB:AL023843).

The 10 amino acid N-terminal domain of human STLK5 does not reveal any significant homologies to the protein database.

15 The 311 amino acid catalytic domain of human STLK5 shares 51% and 34 % identity to STLK6 and SPAK, respectively. The catalytic domain of STLK5 contains a 45 amino acid insert between kinase subdomains X and XI relative to human STE20. Multiple human EST fragments as well as a murine EST (GB:AA575647) contain this insert
20 providing evidence that this region is an integral part of STLK5.

25 The 52 amino acid C-terminal tail of human STLK5 shares 41.3% amino acid identity to human SOK-1 (GB:X99325). The weak sequence homology between the C-termini of human STLK5 and STE20, suggests it may also function as an inhibitory domain for its kinase.

Mammalian STLK6

30 The 2,001 bp human STLK6 nucleotide sequence of the complete cDNA encodes a polypeptide of 418 amino acids (SEQ ID NO:99) with a predicted molecular mass of 47,025 daltons.

Analysis of the deduced amino acid sequence predicts STLK6 to be an intracellular STE20-subfamily kinase, lacking both

a signal sequence and transmembrane domain. STLK6 contains a 57 amino acid N-terminal domain, a 312 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, followed by a 49 amino acid C-terminal domain.

5 STLK6 is most closely related to human STE20-subfamily kinases STLK5 (SEQ ID NO:97), STLK7 (SEQ ID NO:101), and SPAK (AFO99989), sharing 50%, 35%, and 30% amino acid identity over its full extent. It also shares significant 10 homology to database entries from *Arabidopsis thaliana* (GB:AC002343) and *C.elegans* (GB:U53153).

The 57 amino acid N-terminal domain of human STLK6 does not reveal any significant homologies in the protein database.

15 The 312 amino acid catalytic domain of human STLK6 shares 51 and 30 % identity to human STLK5 and SPAK, respectively.

The 49 amino acid C-terminal tail of human STLK6 shares low amino acid sequence identity (29%) with STLK5 and SPAK.

20 Mammalian STLK7
The 311 bp human STLK7 nucleotide sequence of the partial cDNA encodes a polypeptide of 103 amino acids (SEQ ID NO:101). Analysis of the deduced amino acid sequence 25 predicts STLK7 to be an internal fragment of an intracellular STE20-family kinase. This sequence lacks the N- and C-terminal portions of STLK7 and contains only the N-terminal 103 amino acids of the predicted catalytic domain.

30 Human STLK7 is most closely related to human STE20-subfamily kinases SPAK (AFO99989), STLK5 (SEQ ID NO:97), and STLK6 (SEQ ID NO:99), sharing 86%, 38%, and 35% amino acid identity within this region of the kinase domain. It also

shares significant homology to database entries from *Arabidopsis thaliana* (GB:AC002343) and *Drosophila melanogaster* (GB:AF006640).

5 Mammalian ZC1

The 3798 bp human ZC1 nucleotide sequence encodes a polypeptide of 1239 amino acids (SEQ ID NO:13) with a predicted molecular mass of 142,140 daltons. Analysis of the deduced amino acid sequence predicts ZC1 to be an 10 intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. The full-length ZC1 protein contains a 22 amino acid N-terminus, a 267 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, a 237 amino acid region predicted 15 to form a coiled-coil structure, a 114 amino acid proline-rich region, a 256 amino acid spacer region, followed by a 343 amino acid C-terminal domain containing a potential Rab/Rho-binding region.

ZC1 is most closely related to the human STE20-20 subfamily kinases ZC2 (SEQ ID NO:14), ZC3 (SEQ ID NO:15), and ZC4 (SEQ ID NO:16), sharing 61.7%, 60.9%, and 43.8% amino acid identity, respectively. ZC1 also shares 45.5% amino acid identity to a *C. elegans* kinase encoded by the cosmid ZC504.4 (GB:Z50029). ZC1 exhibits 90.0% amino acid 25 homology to murine NIK (GB:U88984), suggesting it may be the human orthologue of this STK.

The 22 amino acid N-terminal domain of human ZC1 is 58.8% identical to the *C. elegans* kinase encoded by the 30 cosmid ZC504.4 (GB:Z50029), and 100% identical to murine NIK (GB: U88984). Human ZC1 lacks a glycine residue at position 2, and is therefore unlikely to undergo myristylation. A Smith-Waterman search of the nonredundant protein database

does not reveal any significant homologies that might suggest a potential function for this domain.

The 267 amino acid catalytic domain of human ZC1 is most related to human STE20-subfamily kinases, ZC3 (SEQ ID NO:15), ZC2 (SEQ ID NO:14), KHS2 (SEQ ID NO:18), SOK-1 (GB:X99325), GCK (GB:U07349), and GEK2 (SEQ ID NO:107), and to the *C. elegans* kinase encoded by the cosmid ZC504.4 (GB:Z50029) sharing 90.6%, 90.2%, 50.6%, 47.4%, 45.4%, 42.5% and 82.6% amino acid identity, respectively. The ZC1 kinase domain shares 98.1% identity to murine NIK (GB:U88984). ZC1 contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC2, ZC3, ZC4, GEK2, KHS2, SULU1, SULU3, PAK4 and PAK5.

Immediately C-terminal to the kinase domain of human ZC1 is a 237 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm (Lupas, A. Meth. Enzymol. 266, 513-525 (1996)). This region of ZC1 is most related to human STE20-subfamily kinases, ZC3 (SEQ ID NO:15), ZC2 (SEQ ID NO:14), and GEK2 (SEQ ID NO:107), as well as to human PITSLRE (GB:U04824) sharing 65.5%, 65.4%, 25.3%, and 29.0% amino acid identity, respectively. The ZC1 coiled-coil domain also shares 90.6% amino acid homology to murine NIK. The *C. elegans* homologue ZC504.4 shares 32.2% sequence identity over this region.

Within the predicted coiled-coil domain of human ZC1, and the related ZC3, is a region predicted to form a leucine zipper (Leu-X6-Leu-X6-Leu-X6-Leu-X20-Leu-X6-Leu). The fact that this leucine repeat exists within a predicted coiled-coil structure suggests that the leucine zipper may have a high probability of serving as a dimerization interface (Hirst, J.D. et al Protein Engineering 9 657-662 (1996))

mediating a potential inter- or intra-molecular dimerization of human ZC1.

The 114 amino acid proline-rich region of human ZC1 is most related to human STE20-subfamily kinases, ZC2 (SEQ ID NO:14) and ZC3 (SEQ ID NO:15), sharing 35.8%, and 24.9%, respectively. The ZC1 proline-rich domain shares 36.4% amino acid homology to murine NIK (GB:U88984). Three potential "PxxP" SH3 domain-binding motifs (I, II and III) are found within the proline-rich region of human ZC1. Motif I is conserved in human ZC1 and *C. elegans* ZC504.4 (GB:Z50029). Motif II is conserved in ZC1, ZC2, ZC3, ZC4 and *C. elegans* ZC504.4. Motif III is conserved in ZC1, ZC2, ZC3 and ZC4. Motifs II and III of murine NIK have been shown to bind the SH3 motif of the adaptor molecule Nck (Su, Y-C. et al, EMBO J. 16, 1279-1290 (1997)). From this evidence, human ZC1 may have the potential to bind to Nck or other SH3 or WW domain-containing proteins and participate in growth factor-induced signaling pathways.

The 256 amino acid spacer region of human ZC1 is most related to human STE20-subfamily kinases, ZC2 (SEQ ID NO:14) and ZC3 (SEQ ID NO:15), as well as to human PITSLRE (GB:U04824), sharing 59.9%, 33.1%, 29.6%, and 26.4% amino acid identity, respectively. It also shares 59.9% amino acid homology to murine NIK. The *C. elegans* homologue ZC504.4 has only limited sequence similarity in this spacer region.

The 343 amino acid C-terminal of human ZC1 is most related to human STE20-subfamily kinases, ZC3 (SEQ ID NO:15), ZC2 (SEQ ID NO:14), and ZC4 (SEQ ID NO:16), sharing 89.2%, 88.9%, and 42.3%, amino acid identity, respectively.

The ZC1 C-terminal domain also shares 98.8% amino acid identity to murine NIK. The *C. elegans* homologue ZC504.4 also shares 68.7% amino acid identity with the C-tail of

human ZC1. A lower, yet significant, homology is also evident to human KHS2 (SEQ ID NO:18), GCK (GB:U07349), and murine citron (GB:U07349) with 26.6%, 23.1% and 36.2% amino acid identity, respectively. GCK is a STE20-family kinase whose C-terminal domain has been shown to bind the small G-protein Rab8 (Ren, M. et al., Proc. Natl. Acad. Sci. 93, 5151-5155 (1996)). Citron is a non-kinase Rho-binding protein (Madaule, P. et al., FEBS Lett. 377, 243-238 (1995)).

The sequence similarity of the C-terminal region of ZC1 to proteins that have potential Rab- or Rho-binding domains suggests that ZC1 may signal through a small G-protein-dependant pathway.

15 Mammalian ZC2

The 4055 bp human ZC2 nucleotide sequence of the partial cDNA encodes a polypeptide of 1297 amino acids (SEQ ID NO:14) with a predicted molecular mass of 147,785 daltons. Analysis of the deduced amino acid sequence predicts ZC2 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain, however the cDNA clone lacks an initiating ATG, so the full extent of its N-terminus is not known. The N-terminally truncated ZC2 protein contains a 255 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, a 187 amino acid region predicted to form a coiled-coil structure, a 184 amino acid proline-rich region, a 328 amino acid spacer region, followed by a 343 amino acid C-terminal domain containing a potential Rab/Rho-binding region.

ZC2 is most closely related to the human STE20-subfamily kinases ZC3 (SEQ ID NO:15), ZC1 (SEQ ID NO:13), and ZC4 (SEQ ID NO:16), sharing 88.3%, 61.7%, and 41.9%

amino acid identity, respectively, and shares 41.7% amino acid identity to a *C. elegans* kinase encoded by the cosmid ZC504.4 (GB:Z50029).

5 The 255 amino acid catalytic domain of human ZC2 is most related to human STE20-subfamily kinases, ZC1 (SEQ ID NO:13), ZC3 (SEQ ID NO:15), SOK-1 (GB:X99325), KHS2 (SEQ ID NO:18), MST1 (GB:U18297), and GCK (GB:U07349), and to the *C. elegans* kinase encoded by the cosmid ZC504.4 (GB:Z50029) sharing .90.2%, 89.8%, 49.0%, 48.6%, 47.9%, 45.0 and 76.7%
10 amino acid identity, respectively. ZC2 contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC1, ZC3, ZC4, GEK2, KHS2, SULU1, SULU3, PAK4 and PAK5.

15 Immediately C-terminal to the kinase domain of human ZC2 is a 187 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm (supra). This region of ZC2 is most related to human STE20-subfamily kinases, ZC1 (SEQ ID NO:13), ZC3 (SEQ ID NO:15), and GEK2 (SEQ ID NO:107), as well as to human PITSLRE (GB:U04824), sharing 65.8%, 61.5%, 29.7% and 29.6% amino acid identity, respectively. The *C. elegans* homologue ZC504.4 shares 30.8% sequence identity over this region. Human ZC2 lacks the potential leucine zipper found in ZC1 as a consequence of a 29 amino acid deletion relative to ZC1 and ZC3.
25

30 The 184 amino acid proline-rich region of human ZC2 is most related to human STE20-subfamily kinases, ZC3 (SEQ ID NO:15) and ZC1 (SEQ ID NO:13), sharing 35.9% and 28.6%, amino acid identity, respectively. Significant homology is also evident to the murine WW domain-binding protein WBP7 (GB:U92455), and to the human SH3 domain-binding protein 3BP-1 (GB:X87671), with 27.7% and 25.3% amino acid identity, respectively.

ZC2 contains two of the potential "PxxP" SH3 domain-binding motifs (II and III) found within the proline-rich region of human ZC1. Motif II is conserved in ZC1, ZC3, ZC4 and *C. elegans* ZC504.4, and Motif III is conserved in ZC1, ZC3 and ZC4. Motifs II and III of murine NIK have been shown to bind the SH3 motif of the adaptor molecule Nck. From this evidence, human ZC1 may have the potential to bind to Nck or other SH3 or WW domain-containing proteins, and to participate in growth factor-induced signaling pathways.

The 328 amino acid spacer region of human ZC2 is most related to human STE20-subfamily kinases ZC1 (SEQ ID NO:13) and ZC3 (SEQ ID NO:15), and to murine NIK (GB:U88984), sharing 31.6%, 26.9% and 25.9% amino acid identity, respectively. The *C. elegans* homologue ZC504.4 has only limited sequence similarity in this spacer region.

The 343 amino acid C-terminal of human ZC2 is most related to human STE20-subfamily kinases ZC1 (SEQ ID NO:13), ZC3 (SEQ ID NO:15) and ZC4 (SEQ ID NO:16), and to murine NIK (GB:U88984), sharing 88.9%, 88.3%, 41.9%, and 88.0%, amino acid identity, respectively. The *C. elegans* homologue, ZC504.4, also shares 67.2% amino acid identity with the C-tail of human ZC2. A lower, yet significant, homology is also evident to human GCK (GB:U07349), murine citron (GB:U07349), and the *S. cerevisiae* ROM2 protein (GB:U19103), a Rho1 GDP/GTP exchange factor, with 22.3%, 22.2% and 21.9% amino acid identity, respectively.

The sequence similarity of the C-terminal region of ZC2 to proteins that have potential Rab- or Rho-binding domains suggests that ZC2, like ZC1, may also signal through a small G-protein-dependant pathway.

Mammalian ZC3

The 4133 bp human ZC3 nucleotide sequence of the partial cDNA encodes a polypeptide of 1326 amino acids (SEQ ID NO:15) with a predicted molecular mass of 149,906 daltons. Analysis of the deduced amino acid sequence predicts ZC3 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain, however the cDNA clone lacks an initiating ATG, so the full extent of its N-terminus is not known. The N-terminally truncated ZC3 protein contains a 255 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase: a 221 amino acid region predicted to form a coiled-coil structure, a 204 amino acid proline-rich region, and a 303 amino acid spacer region followed by a 343 amino acid C-terminal domain containing a potential Rab/Rho-binding region.

ZC3 is most closely related to the human STE20-subfamily kinases ZC1 (SEQ ID NO:13), ZC2 (SEQ ID NO:14), and ZC4 (SEQ ID NO:16), sharing 62.0%, 61.0%, and 42.5% amino acid identity, respectively and shares 46.7% amino acid identity to a *C. elegans* kinase encoded by the cosmid ZC504.4 (GB:Z50029).

The 255 amino acid catalytic domain of human ZC3 is most related to human STE20-subfamily kinases, ZC1 (SEQ ID NO:13), ZC2 (SEQ ID NO:14), SOK-1 (GB:X99325), KHS2 (SEQ ID NO:18), GCK (GB:U07349), SULU1 (SEQ ID NO:22), and GEK2 (SEQ ID NO:107), and to the *C. elegans* kinase encoded by the cosmid ZC504.4 (GB:Z50029) sharing 90.6%, 89.3%, 49.0%, 48.3%, 45.0%, 43.1%, 42.3% and 76.7% amino acid identity, respectively. ZC1 contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC1, ZC2, GEK2, KHS2, SULU1, SULU3, PAK4 and PAK5.

Immediately C-terminal to the kinase domain of human ZC3 is a 221 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm (supra). This region of ZC3 is most homologous to human STE20-subfamily kinases, ZC1 (SEQ ID NO:13), ZC2 (SEQ ID NO:14), and GEK2 (SEQ ID NO:107), sharing 66.9%, 61.5%, and 27.5% identity, as well as to rat PLC-beta (GB:A45493) and human PITSLRE (GB:H54024) sharing 29.6% and 25.9% amino acid identity, respectively. The *C. elegans* homologue ZC504.4 shares 26.8% sequence identity over this region.

Within the predicted coiled-coil domain of human ZC3, and the related ZC1, is a region predicted to form a leucine zipper (Leu-X6-Leu-X6-Leu-X6-Leu-X20-Leu-X6-Leu). The fact that this leucine repeat exists within a predicted coiled-coil structure suggests that the leucine zipper may have a high probability of serving as a dimerization interface (Hirst, J.D. et al Protein Engineering 9 657-662 (1996)) mediating a potential inter- or intra-molecular dimerization of human ZC3.

The 204 amino acid proline-rich region of human ZC3 is most related to human STE20-subfamily kinases, ZC1 (SEQ ID NO:13) and ZC2 (SEQ ID NO:14), sharing 66.9% and 61.5% amino acid identity, respectively.

ZC3 contains two of the potential "PxxP" SH3 domain-binding motifs (II and III) found within the proline-rich region of human ZC1. Motif II is conserved in ZC1, ZC2, ZC4 and *C. elegans* ZC504.4; Motif III is conserved in ZC1, ZC2 and ZC4. Motifs II and III of murine NIK have been shown to bind the SH3 motif of the adaptor molecule Nck. From this evidence, human ZC3 may have the potential to bind to Nck or other SH3 or WW domain-containing proteins and participate in growth factor-induced signaling pathways.

The 303 amino acid acid spacer region of human ZC3 is most related to human STE20-subfamily kinases, ZC1 (SEQ ID NO:13) and ZC2 (SEQ ID NO:14) sharing 30.1%, and 27.1% amino acid identity, respectively. The *C. elegans* homologue ZC504.4 lacks nearly the entire spacer region of ZC3.

5 The 343 amino acid C-terminal of human ZC3 is most related to human STE20-subfamily kinases, ZC1 (SEQ ID NO:13), ZC2 (SEQ ID NO:14) and ZC4 (SEQ ID NO:16), sharing 89.2%, 88.9%, and 42.5%, amino acid identity, respectively.
10 The *C. elegans* homologue ZC504.4 also shares 67.2% amino acid identity with the C-tail of human ZC3. A lower, yet significant, homology is also evident to human GCK (GB:U07349), as well as to the non-kinases murine citron (GB:U07349) and the *S. cerevisiae* ROM2 protein (GB:U19103),
15 a Rho1 GDP/GTP exchange factor, with 21.6%, 32.4% and 22.9% amino acid identity, respectively.

20 The sequence similarity of the C-terminal region of ZC3 to proteins that have potential Rab- or Rho-binding domains suggests that ZC3, like ZC1 and ZC2, may signal through a small G-protein-dependant pathway.

Mammalian ZC4

25 The 3,684 bp human ZC4 nucleotide sequence of the complete cDNA encodes a polypeptide of 1,227 amino acids (SEQ ID NO:105) with a predicted molecular mass of 138,205 Daltons. Analysis of the deduced amino acid sequence predicts ZC4 to be an intracellular STE20-subfamily kinase, lacking both a signal sequence and a transmembrane domain.
30 The full-length ZC4 protein contains a 15 amino acid N-terminus, a 265 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, a 108 amino acid region predicted to form a cciled-coil structure, a 231 amino acid proline-rich region, a 40 amino acid region

predicted to form a coiled-coil structure spacer region, a 204 amino acid spacer region (domain B), followed by a 355 amino acid C-terminal domain containing a potential Rab/Rho-binding region (domain C).

5 ZC4 is most closely related to human ZC1 (SEQ ID NO:13, also known as human HGK, human KIAA0687, murine NIK, human AC005035, human NIK, and *C. elegans* MIG-15), ZC2 (SEQ ID NO:14, similar to partial sequence human KIAA0551), and ZC3 (SEQ ID NO:15). An assembled genomic fragment in the
10 database (Z83850) is identical to ZC4, except for inappropriate identification of the exon boundaries. (Abo et al. (1998) EMBO J. 17: 6527-6540.)

15 The 25 amino acid N-terminal domain of human ZC4 shares weak homology to human ZC1 in its C-terminal extent, but otherwise does not reveal any significant homologies to the protein database.

20 The 265 amino acid catalytic domain of human ZC4 is most related to human ZC1 (SEQ ID NO:13), ZC3 (SEQ ID NO:15), and ZC2 (SEQ ID NO:14), sharing 63%, 64% and 62% amino acid identity, respectively.

25 Immediately C-terminal to the kinase domain of human ZC4 is a 108 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm (supra). This region is most related to human ZC1 (SEQ ID NO:13), ZC3 (SEQ ID NO:15), and ZC2 (SEQ ID NO:14), sharing 29%, 25% and 20% amino acid identity, respectively.

30 The 231 amino acid proline-rich region of human ZC4 does not reveal any significant homologies to the protein database. This region of ZC4 contains two "PxxP" motifs that could potentially bind to proteins containing SH3 or WW domains and allow ZC4 to participate in growth factor activated signaling pathways. In addition, within the pro-rich domain of human ZC4 is a region predicted to form a

leucine zipper (Leu-X6-Leu-X6-Leu-X6-Leu-X20-Leu-X6-Leu) which may serve as a dimerization interface. The ZC STE20 subfamily kinases (ZC1, ZC2 and ZC3) have similarly located "PxxP" motifs and potential Leu zippers.

5 Immediately C-terminal to the proline-rich region of human ZC4 is a 40 amino acid region also predicted to form a coiled-coil structure based on the Lupas algorithm. This region of human ZC4 does not reveal any significant homologies to the protein database.

10 The 204 amino acid acidic- and serine-rich domain "B" of ZC4 does not reveal any significant homologies to the protein database.

15 The 355 amino acid C-terminal of human ZC4 is most related to human ZC1 (SEQ ID NO:13), ZC3 (SEQ ID NO:15), and ZC2 (SEQ ID NO:14), sharing 43%, 42% and 42% amino acid identity, respectively.

20 The sequence similarity of the C-terminal region of ZC4 to proteins that have potential Rab- or Rho-binding domains suggests that ZC4, like other ZC-subfamily STE20 kinases, may signal through a small G-protein-dependant pathway.

Mammalian KHS2

25 The 4023 bp human KHS2 nucleotide sequence encodes a polypeptide of 894 amino acids (SEQ ID NO:18) with a predicted molecular mass of 101,327 daltons. Analysis of the deduced amino acid sequence predicts KHS2 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. The full-length KHS2 protein contains a 13 amino acid N-terminus, a 260 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, a 73 amino acid spacer region, a 188 proline-rich region, followed by a 360 amino acid C-terminal domain containing a potential Rab/Rho-binding site.

KHS2 is most closely related to the human STE20-subfamily kinases KHS1 (GB:U177129), GCK (GB:U07349), and HPK1 (GB:U07349), sharing 65.5%, 51.9%, and 44.9% amino acid identity, respectively and shares 38.5% amino acid identity to a *C. elegans* STK (GB:U55363).

The 13 amino acid N-terminal domain of human KHS2 does not reveal any significant homologies that might suggest a potential function for this domain when examined by a Smith-Waterman alignment to the nonredundant protein database.

10 Human KHS2 lacks a glycine residue at position 2, and is therefore unlikely to undergo myristylation.

The 260 amino acid catalytic domain of human KHS2 is most related to human STE20-subfamily kinases KHS1 (GB:U177129), GCK (GB:U07349), HPK1 (GB:U66464), SOK-1 (GB:X99325), MST1 (GB:U18297), ZC1 (SEQ ID NO:13), and to the *C. elegans* kinase (GB:U55363), sharing 85.4%, 75.1%, 67.7%, 51.4%, 48.1%, 49.8% and 72.0% amino acid identity, respectively. KHS2 contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC1, ZC2, ZC3, ZC4, GEK2, SULU1, SULU3, PAK4 and PAK5.

The 73 amino acid acid spacer region of human KHS2 is most related to human STE20-subfamily kinases, KHS1 (GB:U177129), HPK1 (GB:U66464) and GCK (GB:U07349), sharing 60.3%, 43.5% and 44.0%, amino acid identity, respectively.

The 188 amino acid proline-rich region of human KHS2 is most related to human STE20-subfamily kinases, HPK1 (GB:U66464), GCK (GB:U07349) and KHS1 (GB:U177129), sharing 33.3%, 31.9% and 31.4%, amino acid identity, respectively.

Two potential "PxxP" SH3 domain-binding motifs (I and II) are found within the proline-rich region of human KHS2. Motif I is conserved with human KHS1 and HPK1; motif II is

conserved with GCK and KHS2. A 192 amino acid region of human HPK1 containing motif II has been shown to bind to the C-terminal SH3 motif of the adaptor molecule Grb2 (Anafi, M et al, J. Biol. Chem. J. 272, 27804-27811 (1997)). Human KHS2 may bind SH3 or WW domain-containing proteins through this proline-rich region.

The 360 amino acid C-terminal of human KHS2 is most related to KHS1 (GB:U177129), GCK (GB:U07349) and HPK1 (GB:U66464), and to the *C. elegans* kinase (GB:U55363), sharing 74.9%, 54.8%, 42.9%, and 31.0%, amino acid identity, respectively. GCK is a STE20-family kinase whose C-terminal domain has been shown to bind the small G-protein Rab8 (Ren, M. et al., Proc. Natl. Acad. Sci. 93, 5151-5155 (1996)).

15 Mammalian SULU1

The 4196 bp human SULU1 nucleotide sequence encodes a polypeptide of 898 amino acids (SEQ ID NO:22) with a predicted molecular mass of 105,402 daltons. Analysis of the deduced amino acid sequence predicts SULU1 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. The full-length SULU1 protein contains a 21 amino acid N-terminus, a 256 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, a 150 amino acid spacer region, a 210 amino acid region predicted to form a coiled-coil structure, a 114 amino acid spacer region and a 147 amino acid C-terminal domain predicted to form a coiled-coil structure.

SULU1 is most closely related to the STE20-subfamily kinases murine SULU3 (SEQ ID NO:24), human SULU3 (SEQ ID NO:23), and to the *C. elegans* kinase SULU (GB:U11280), sharing 68.9%, 72.2% and 38.2% amino acid identity, respectively.

The 21 amino acid N-terminal domain of human SULU1 is most related to murine SULU3 (SEQ ID NO:24) and to the *C. elegans* kinase SULU (GB:U11280), sharing 86.3% and 62.3% amino acid identity. Human SULU1 lacks a glycine residue at position 2, and is therefore unlikely to undergo myristoylation. A Smith-Waterman search of the nonredundant protein database does not reveal any significant homologies that might suggest a potential function for this domain.

The 256 amino acid catalytic domain of human SULU1 is most related to murine SULU3 (SEQ ID NO:24), and to human SOK-1 (GB:X99325), STLK2 (SEQ ID NO:5), MST1 (GB:U18297), PAK1 (GB:U24152), ZC2 (SEQ ID NO:14), and KHS2 (SEQ ID NO:18) sharing 86.3%, 48.1%, 46.9%, 45.2%, 43.3%, 43.1% and 42.0% amino acid identity, respectively. The *C. elegans* SULU STK (GB:U11280) shares 62.3% sequence identity over this region. SULU1 contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC1, ZC2, ZC3, ZC4, GEK2, KHS2, SULU3, PAK4 and PAK5.

The 150 amino acid spacer region of human SULU1 is most related to human SULU3 (SEQ ID NO:23) and to the *C. elegans* kinase (GB:U11280), sharing 53.5% and 10.4% amino acid identity, respectively.

Immediately C-terminal to the spacer region of human SULU1 is a 210 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm. This region of SULU1 is most related to SULU3 (SEQ ID NO:23), the *C. elegans* SULU kinase (GB:U11280), GEK 2 (SEQ ID NO:107) and ZC1 (SEQ ID NO:13), sharing 68.6%, 26.8%, 23.2%, and 22.8% amino acid identity, respectively.

The 114 amino acid spacer region human SULU1 is most related to human SULU3 (SEQ ID NO:24) with 73.7% amino acid

sequence identity. A lower, yet significant, homology is also evident to murine PITSLRE (GB:U04824) and DLK (GB:A55318), human ZC1 (SEQ ID NO:13) and GEK 2 (SEQ ID NO:107), as well as to the *C. elegans* SULU STK (GB:U11280), sharing 39.7%, 35.4%, 29.5%, 23.6% and 37.6% amino acid identity, respectively.

Immediately C-terminal to the second spacer region of human SULU1 is a 147 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm. This region of SULU1 is most related to human SULU3 (SEQ ID NO:24), ZC1 (SEQ ID NO:13) and GEK 2 (SEQ ID NO:107), as well as to the *C. elegans* SULU STK (GB:U11280), sharing 73.3%, 28.4%, 26.1% and 39.5%, amino acid identity, respectively.

15

Mammalian (human) SULU3

The 3824 bp partial cDNA human SULU3 nucleotide sequence encodes a polypeptide of 786 amino acids (SEQ ID NO:23) with a predicted molecular mass of 92,037 daltons. Analysis of the deduced amino acid sequence predicts SULU3 to be an intracellular serine/threonine kinase lacking a transmembrane domain. The N-terminally truncated human SULU3 protein contains a 66 amino acid partial catalytic domain followed by a 149 amino acid spacer region, a 210 amino acid region predicted to form a coiled-coil structure, a second spacer region of 114 amino acids, a 247 amino acid C-terminal region predicted to form a second coiled-coil structure and a 100 amino acid C-terminal tail.

Human SULU3 is most closely related murine SULU3 (SEQ ID NO:24), human SULU1 (SEQ ID NO:22), and to the *C. elegans* SULU kinase (GB:U11280), sharing 66.3%, 68.9% and 32.9% amino acid identity, respectively. The high sequence

homology between murine and human SULU3 suggests that these two proteins are orthologs of each other.

The 66 amino acid partial catalytic domain of human SULU3 is most related to murine SULU3 (SEQ ID NO:24), and to the human STE20 subfamily kinases ZC1 (SEQ ID NO:13), STE20 (GB:X99325), KHS1 (GB:U177129) and GEK 2 (SEQ ID NO:107), as well as to the *C. elegans* SULU kinase (GB:U11280), sharing 83.3%, 47.0%, 45.5%, 43.5%, 41.8% and 55.6% amino acid identity, respectively.

The 149 amino acid spacer region of human SULU3 is most related to murine SULU3 (SEQ ID NO:24), human STE20 (GB:X99325), MST1 (GB:U18297), and to the *C. elegans* SULU kinase (GB:U11280) sharing 98.7%, 21.9% and 21.8% amino acid identity, respectively.

Immediately C-terminal to the first spacer region of human SULU3 is a 210 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm. This region of SULU3 is most related to murine SULU3 (SEQ ID NO:24), and to human SULU1 (SEQ ID NO:22), ZC1 (SEQ ID NO:13) and GEK 2 (SEQ ID NO:107), as well as to the *C. elegans* SULU kinase (GB:U11280), sharing 99.5%, 68.6%, 27.4% and 22.5% amino acid identity, respectively.

The 114 amino acid second spacer region of human SULU3 is most related to murine SULU3 (SEQ ID NO:24), and to human SULU1 (SEQ ID NO:22) GEK 2 (SEQ ID NO:107), and ZC1 (SEQ ID NO:13), as well as to the *C. elegans* SULU kinase (GB:U11280), sharing 99.1%, 73.7%, 24.6%, 24.1% and 41.2% amino acid identity, respectively.

Immediately C-terminal to the second spacer region of human SULU3 is a 247 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm (supra).

This region of SULU3 is most related to human SULU1 (SEQ ID NO:22) and ZC1 (SEQ ID NO:13) as well as to rat PKN-

(GB:D26180) murine p160 ROCK1 (GB:U58512), and the *C. elegans* SULU kinase (GB:U11280), sharing 73.7%, 26.7%, 24.0% and 21.0% amino acid identity, respectively.

The 100 amino acid C-tail of human SULU3 is most related to a human prion protein (GB:L38993), with 45.0% amino acid identity.

Mammalian (murine) SULU3

The 2249 bp murine, partial cDNA SULU3 nucleotide sequence encodes a polypeptide of 748 amino acids (SEQ ID NO:24) with a predicted molecular mass of 87,520 daltons. Analysis of the deduced amino acid sequence predicts SULU3 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. The partial murine SULU3 protein contains a 25 amino acid N-terminus, a 248 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, a 149 amino acid spacer region, a 210 amino acid region predicted to form a coiled-coil structure, and a 116 amino acid spacer region.

Murine SULU3 is most closely related to human SULU3 (SEQ ID NO:23) and SULU1 (SEQ ID NO:22), as well as to the *C. elegans* SULU kinase (GB:U11280), sharing 97.0%, 72.3% and 38.4% amino acid identity, respectively. The high sequence homology between murine and human SULU3 suggests that these two proteins are orthologs.

The 25 amino acid N-terminal domain of murine SULU3 is most related to human SULU1 (SEQ ID NO:22) and to the *C. elegans* SULU kinase (GB:U11280), sharing 70.0% and 44.4% amino acid identity, respectively.

Murine SULU3 lacks a glycine residue at position 2, and is therefore unlikely to undergo myristoylation. A Smith-Waterman search of the nonredundant protein database does

not reveal any significant homologies that might suggest a potential function for this domain.

The 248 amino acid catalytic domain of murine SULU3 is most related to human SULU1 (SEQ ID NO:22), STE20 (GB:X99325), ZC1 (SEQ ID NO:13), and KHS1 (GB:U77129), as well as to the *C. elegans* SULU kinase (GB:U11280), sharing 86.7%, 46.6%, 43.3%, 59.4% amino acid identity, respectively. Murine SULU3 contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC2, ZC3, ZC4, GEK2, KHS2, SULU1, SULU3, PAK4 and PAK5.

The 149 amino acid spacer of murine SULU3 is most related to human SULU3 (SEQ ID NO:23), SULU1 (SEQ ID NO:22), and STE20 (GB:X99325), as well as to the *C. elegans* SULU (GB:U11280) and the *S. cerevisiae* STE20 (GB:L04655) kinases, sharing 98.7%, 53.4%, 21.9%, 59.4% and 21.9% amino acid identity, respectively.

Immediately C-terminal to the spacer region of murine SULU3 is a 210 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm. This region of murine SULU3 is most related to human SULU3 (SEQ ID NO:23), ZC1 (SEQ ID NO:13), and GEK 2 (SEQ ID NO:107), as well as to the *C. elegans* SULU kinase (GB:U11280), sharing 99.5%, 27.4%, 22.5% and 29.2% amino acid identity, respectively.

The 116 amino acid C-terminal spacer region of murine SULU3 is most related to human SULU3 (SEQ ID NO:23), GEK 2 (SEQ ID NO:107), and ZC1 (SEQ ID NO:13), well as to the *C. elegans* SULU kinase (GB:U11280), sharing 98.3%, 24.6%, 24.1% and 40.5% amino acid identity, respectively.

Mammalian (murine/human) SULU3

The 2249 bp murine SULU3 and the 3824 bp human SULU3 cDNAs contain a 1620 nucleotide overlap (541 amino acids) with 90% and 98% DNA and amino acid sequence identity, respectively. Owing to the high degree of sequence identity in this extended overlap, we propose that these are functional orthologues of a single gene. The combined murine/human 4492 bp SULU3 sequence encodes a polypeptide of 1001 amino acids (SEQ ID NO:31) with a predicted molecular mass of 116,069 daltons. Analysis of the deduced amino acid sequence predicts SULU3 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. SULU3 contains a 25 amino acid N-terminus, a 248 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, a 149 amino acid spacer region, a 210 amino acid region predicted to form a coiled-coil structure and a second spacer region of 114 amino acids, a 247 amino acid C-terminal region predicted to form a second coiled-coil structure and a 100 amino acid C-terminal tail. The murine SULU3 clone lacks the region from the second C-terminal coiled-coil to the C-terminus, whereas the human clone lacks the N-terminal domain, and all but 66 amino acids of the 248 amino acid kinase domain.

SULU3 is most closely related to SULU1 (SEQ ID NO:22) and the *C. elegans* SULU kinase (GB:U11280) sharing 72.3% and 38.4% amino acid identity, respectively.

The 25 amino acid N-terminal domain of SULU3 is most related to human SULU1 (SEQ ID NO:22) and to the *C. elegans* SULU kinase (GB:U11280), sharing 70.0% and 44.4% amino acid identity, respectively. SULU3 lacks a glycine residue at position 2, and is therefore unlikely to undergo myristylation. A Smith-Waterman search of the nonredundant

protein database does not reveal any significant homologies that might suggest a potential function for this domain.

The 248 amino acid catalytic domain of SULU3 is most related to human SULU1 (SEQ ID NO:22), SOK-1 (GB:X99325), ZC1 (SEQ ID NO:13), KHS1 (GB:U77129) and the *C. elegans* SULU kinase (GB:U11280), sharing 86.7%, 46.6%, 43.3%, 42.0% and 59.4% amino acid identity, respectively. SULU3 contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC2, ZC3, ZC4, GEK2, KHS2, SULU1, PAK4 and PAK5.

The 149 amino acid spacer of SULU3 is most related to SULU1 (SEQ ID NO:22) and SOK-1 (GB:X99325), and to the *C. elegans* SULU (GB:U11280), and *S. cerevisiae* STE20 (GB:L04655) kinases, sharing 53.4%, 21.9%, 59.4% and 21.9% amino acid identity, respectively.

Immediately C-terminal to the spacer region of SULU3 is a 210 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm. This region is most related to ZC1 (SEQ ID NO:13), GEK 2 (SEQ ID NO:107), and the *C. elegans* SULU kinase (GB:U11280), sharing 27.4% 22.5% and 29.2% amino acid identity, respectively.

The 114 amino acid spacer region of SULU3 is most related to human SULU1 (SEQ ID NO:22), GEK 2 (SEQ ID NO:107), ZC1 (SEQ ID NO:13), and to the *C. elegans* SULU kinase (GB:U11280), sharing 73.7%, 24.6%, 24.1% and 41.2% amino acid identity, respectively.

Immediately C-terminal to the second spacer region of SULU3 is a 247 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm. This region of SULU3 is most related to human SULU1 (SEQ ID NO:22) and ZC1 (SEQ ID NO:13), as well as to rat PKN (GB:D26180), murine p160 ROCK1 (GB:U58512) and the *C. elegans* SULU kinase

(GB:U11280), sharing 73.7%, 26.7%, 24.0%, 21.0% and 37.6% amino acid identity, respectively.

The 100 amino acid C-tail of SULU3 is most related to a human prion protein (GB:L38993) with 45.0% amino acid
5 identity.

Mammalian GEK2

The 2926 bp human GEK2 nucleotide sequence of the complete cDNA encodes a polypeptide of 968 amino acids (SEQ ID NO:107) with a predicted molecular mass of 112,120 daltons. Analysis of the deduced amino acid sequence predicts GEK2 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. The complete GEK2 protein contains a 33 amino acid N-terminus, a 261 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase, a 43 amino acid spacer region, a 135 amino acid proline-rich region, a 252 amino acid region predicted to form a coiled-coil structure followed by a 244 amino acid region also predicted to form a coiled-coil structure.
10
15
20

GEK2 is most closely related to rat AT1-46 (GB:U33472) (a partial cDNA that extends from the middle of the first potential coiled-coil domain of GEK2 to the C-terminus), murine LOK (GB:D89728), *Xenopus laevis* polo-like kinase 1 (GB:AF100165), and human SLK (GB:AB002804), sharing 91.3%, 88.5%, 65.0%, and 44.7% amino acid identity, respectively. The high sequence homology between human GEK2, murine LOK and rat AT1-46 suggests that human GEK1 is a highly related protein to the rodent forms, or alternatively, its orthologue. Recently, a full-length version of GEK2 was reported (STK10 or human LOK AB015718). The 968 amino acid sequence is 99% identical to GEK2 (SEQ ID NO:107).
25
30

The 33 amino acid N-terminal domain of human GEK2 is most related to murine LOK (GB:D89728) and to human SLK (GB:AB002804), sharing 100% and 54.5% amino acid identity, respectively.

5 Human GEK2 lacks a glycine residue at position 2, and is therefore unlikely to undergo myristylation. A Smith-Waterman search of the nonredundant protein database does not reveal any significant homologies that might suggest a potential function for this domain.

10 The 261 amino acid catalytic domain of human GEK2 is most related to murine LOK (GB:D89728), rat AT1-46 (GB:D89728) and human SLK (GB:AB002804) as well as to a *C. elegans* kinase (GB:Z81460), sharing 97.7%, 90.8%, 54.5% and 55.9% amino acid identity, respectively. GEK2 contains the 15 potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC2, ZC3, ZC4, GEK2, KHS2, SULU1, SULU3, PAK4 and PAK5.

20 The 43 amino acid spacer region of human GEK2 is most related to murine LOK (GB:D89728) and to human SLK, sharing 83.7% and 77.6% amino acid identity, respectively.

25 The 135 amino acid proline-rich region of human GEK2 is most related to murine LOK (GB:D89728) with 66.2% amino acid identity, respectively. Within the proline-rich region of human GEK2 is a potential "PxxP" SH3-binding domain, conserved with murine LOK.

30 Immediately C-terminal to the proline-rich region of human GEK2 is a 252 amino acid region predicted to form a coiled-coil structure based on the Lupas algorithm. This region of human GEK2 is most related to rat AT1-46 (GB:D89728), murine LOK (GB:D89728) and human SLK (GB:AB002804), and ZC2 (SEQ ID NO:14), sharing 90.8%, 86.9%, 42.2%, and 29.7% amino acid identity, respectively.

Immediately C-terminal to the predicted coiled-coil structure of human GEK2 is a second potential coiled-coil structure of 244 amino acids predicted based on the Lupas algorithm. This region of human GEK2 is most related to rat AT1-46 (GB:D89728) and murine LOK (GB:D89728) as well as to human SLK (GB:AB002804) and ZC1 (SEQ ID NO:13), sharing 91.8%, 92.6%, 70.4% and 26.7% amino acid identity, respectively. The *C. elegans* kinase (GB:Z81460) shares 31.5% amino acid sequence identity over this region.

10

Mammalian PAK4

The 3604 bp human PAK4 nucleotide sequence encodes a polypeptide of 681 amino acids (SEQ ID NO:29) with a predicted molecular mass of 74,875 daltons. Analysis of the deduced amino acid sequence predicts PAK4 to be an intracellular serine/threonine kinase, lacking both a signal sequence and transmembrane domain. The full-length PAK4 protein contains a 51 amino acid N-terminus predicted to contain a rac-binding motif, a 173 amino acid insert relative to the known mammalian PAK proteins, a 169 amino acid spacer region, a 265 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase and a 23 amino acid C-terminal tail.

PAK4 is most closely related to human PAK5 (SEQ ID NO:30), PAK1 (GB: U24152), and PAK65 (GB:U25975), as well as to a *C. elegans* kinase (GB: Z74029), sharing 76.8%, 49.5%, 49.8%, and 34.6% amino acid identity, respectively.

The 51 amino acid N-terminal domain of human PAK4 is most related to human PAK1 (GB:U24152), and PAK65 (GB:U25975), as well as to a *C. elegans* kinase (GB: Z74029), sharing 50.0%, 50.0% and 49.0% amino acid identity, respectively.

The 11 amino acid region at positions 13-23 of human PAK4 fits the consensus for a Cdc42/Rac-binding motif (SXPX4-6HXXH) (Burbelo, P.D., Dreschel, D. and Hall, A. J. Bio. Chem. 270, 29071-29074 (1995)).

5 The 173 amino acid insert of human PAK4, relative to the known mammalian PAK proteins, is most related to a *C. elegans* kinase (GB: Z74029) with 39.0% amino acid identity.

A Smith-Waterman search of the nonredundant protein database does not reveal any significant homologies that 10 might suggest a potential function for this region.

The 169 amino acid spacer of human PAK4 does not reveal any significant homologies that might suggest a potential function for this region.

15 The equivalent spacer region in PAK1 binds to the guanine nucleotide exchange factor PIX (Manser, E. et al (1998) Molecular Cell, 1, 183-192). Since PAK4 differs substantially from PAK1 over this region, the spacer domain of PAK4 may differ in its guanine nucleotide exchange factor binding specificity, relative to PAK1.

20 The 265 amino acid catalytic domain of human PAK4 is most related to human PAK5 (SEQ ID NO:30), PAK1 (GB:U24152), GCK (GB:U07349), SOK-1 (GB:X99325), and SLK (GB:AB002804), as well as to the *C. elegans* (GB: Z74029), and *S. cerevisiae* STE20-related kinases (GB:L04655), sharing 95.9%, 51.7%, 25 41.3%, 39.8%, 37.4%, 60.2% and 47.9% amino acid identity, respectively. PAK4 contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC1, ZC2, ZC3, ZC4, GEK2, KHS2, SULU1, SULU3 and PAK5.

30 The 23 amino acid C-tail of human PAK4 contains a sequence that is homologous to a G-protein beta subunit binding site (Leeuw, T. et al. Nature, 391, 191-195 (1998)).

PAK4 has, therefore, the potential to be activated by both Cdc42- as well as G-protein-dependant pathways.

Mammalian PAK5

5 The 2,806 bp human PAK5 nucleotide sequence of the complete cDNA encodes a polypeptide of 591 amino acids (SEQ ID NO:103) with a predicted molecular mass of 64,071 Daltons. Analysis of the deduced amino acid sequence predicts PAK5 to be an intracellular STE20-subfamily kinase, 10 lacking both a signal sequence and transmembrane domain. The full-length PAK5 protein contains a 52 amino acid N-terminus predicted to contain a p21 (small G-protein) binding domain (PDB or CRIB), a 121 amino acid insert relative to the known mammalian PAK proteins, a 134 amino 15 spacer region, a 265 amino acid catalytic domain with all the motifs characteristic of a serine/threonine kinase and a 19 amino acid C-terminal tail.

PAK5 is most closely related to Human PAK4 (SEQ ID NO:29), *Drosophila melanogaster* PAK (also known as "mushroom bodies tiny") (AJ011578), C45B11.1b from *C. elegans* (Z74029), and human PAK3 (Q13177) sharing 48% (327/674 aa), 20 50% (330/651 aa), 43% (234/435 aa excluding gap), and 47% (190/405 aa excluding gap) amino acid identity, respectively. Recently, a full length version of PAK5 was 25 reported (PAK4 AF005046) whose 591 amino acid sequence is identical to PAK5 (SEQ ID NO:103). (Abo, et al. (1998) EMBO J. 17:6527-6540).

The 52 amino acid N-terminal domain of human PAK5 is 30 most related to human PAK4 (SEQ ID NO:29), *Drosophila melanogaster* PAK (AJ011578), C45B11.1b from *C. elegans* (Z74029), and human PAK3 (Q13177), sharing 65%, 57%, 54%, and 53% amino acid identity, respectively.

The 11 amino acid region at positions 12-22 of human PAK5 (FIG. 18) fits the consensus for a small G-protein binding domain (PDB or CRIB) (SXPX4-6HXXH) (Burbelo, P.D., Dreschel, D. and Hall, A. J. Bio. Chem. 270, 29071-29074 (1995), hereby incorporated by reference herein in its entirety including any figures, tables, or drawings.).

The 121 amino acid insert of human PAK5 shares 43% amino acid identity with a similar domain from PAK4 (SEQ ID NO:29), but that is absent from other known PAKs.

The equivalent spacer region in PAK1 binds to the guanine nucleotide exchange factor PIX (Manser, E. et al (1998) Molecular Cell, 1, 183-192 hereby incorporated by reference herein in its entirety including any drawings, figures, or tables.). Since PAK5 differs substantially from PAK1 over this region, the spacer domain of PAK5 may differ in its guanine nucleotide exchange factor binding specificity, relative to PAK1.

The 134 amino acid collagen-like region of human PAK5 shares 34% amino acid identity to pro- α I type collagen from several species and is not present in other known PAKs.

The 265 amino acid catalytic domain of human PAK5 is most related to human PAK4 (SEQ ID NO:29), *Drosophila melanogaster* PAK (AJ011578), C45B11.1b from *C. elegans* (Z74029), and human PAK3 (Q13177), sharing 78%, 80%, 61%, and 55% amino acid identity, respectively. PAK5 also contains the potential "TPY" regulatory phosphorylation site in its activation loop. This "TPY" motif is conserved in other STE20-related kinases, including ZC1, ZC2, ZC3, ZC4, GEK2, KHS2, SULU1, SULU3 and PAK4.

The 19 amino acid C-tail shares 80% amino acid identity to a PAK-like homologue identified from genomic DNA (AL031652). Furthermore, this C-terminal region of human PAK5 contains a sequence that is homologous to a G-protein

beta subunit binding site (Leeuw, T. et al. Nature, 391, 191-195 (1998) hereby incorporated by reference herein in its entirety including any figures, tables, or drawings). PAK5 has, therefore, the potential to be activated by both, 5 Cdc42 as well as G-protein-dependant pathways.

V. Antibodies, Hybridomas, Methods of Use and Kits for Detection of STE20-Related Kinases

The present invention relates to an antibody having 10 binding affinity to a kinase of the invention. The polypeptide may have the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, or a functional derivative thereof, or at least 9 contiguous amino acids thereof (preferably, at least 20, 30, 35, or 40 or more contiguous amino acids thereof).

The present invention also relates to an antibody 20 having specific binding affinity to a kinase of the invention. Such an antibody may be isolated by comparing its binding affinity to a kinase of the invention with its binding affinity to other polypeptides. Those which bind selectively to a kinase of the invention would be chosen for 25 use in methods requiring a distinction between a kinase of the invention and other polypeptides. Such methods could include, but should not be limited to, the analysis of altered kinase expression in tissue containing other polypeptides.

30 The STE20-Related kinases of the present invention can be used in a variety of procedures and methods, such as for the generation of antibodies, for use in identifying

pharmaceutical compositions, and for studying DNA/protein interaction.

The kinases of the present invention can be used to produce antibodies or hybridomas. One skilled in the art will recognize that if an antibody is desired, such a peptide could be generated as described herein and used as an immunogen. The antibodies of the present invention include monoclonal and polyclonal antibodies, as well fragments of these antibodies, and humanized forms.

Humanized forms of the antibodies of the present invention may be generated using one of the procedures known in the art such as chimerization or CDR grafting.

The present invention also relates to a hybridoma which produces the above-described monoclonal antibody, or binding fragment thereof. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

In general, techniques for preparing monoclonal antibodies and hybridomas are well known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands, 1984; St. Groth et al., J. Immunol. Methods 35:1-21, 1980). Any animal (mouse, rabbit, and the like) which is known to produce antibodies can be immunized with the selected polypeptide. Methods for immunization are well known in the art. Such methods include subcutaneous or intraperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of polypeptide used for immunization will vary based on the animal which is immunized, the antigenicity of the polypeptide and the site of injection.

The polypeptide may be modified or administered in an adjuvant in order to increase the peptide antigenicity. Methods of increasing the antigenicity of a polypeptide are well known in the art. Such procedures include coupling the antigen with a heterologous protein (such as globulin or β -galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, such as SP2/0-Agl4 myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells. Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an antibody with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz et al., Exp. Cell Res. 175:109-124, 1988). Hybridomas secreting the desired antibodies are cloned and the class and subclass are determined using procedures known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology", *supra*, 1984).

For polyclonal antibodies, antibody-containing antisera is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures. The above-described antibodies may be detectably labeled. Antibodies can be detectably labeled through the use of radioisotopes, affinity labels (such as biotin, avidin, and the like), enzymatic labels (such as horse radish peroxidase, alkaline phosphatase, and the like) fluorescent labels (such as FITC or rhodamine, and the like), paramagnetic atoms, and the like. Procedures for accomplishing such labeling are well-known in the art, for example, see Stemberger et al., J.

Histochem. Cytochem. 18:315, 1970; Bayer et al., Meth. Enzym. 62:308-, 1979; Engval et al., Immunol. 109:129-, 1972; Goding, J. Immunol._Meth. 13:215-, 1976. The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays to identify cells or tissues which express a specific peptide.

The above-described antibodies may also be immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir et al., "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10, 1986; Jacoby et al., Meth. Enzym. 34, Academic Press, N.Y., 1974). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays as well as in immunochromotography.

Furthermore, one skilled in the art can readily adapt currently available procedures, as well as the techniques, methods and kits disclosed herein with regard to antibodies, to generate peptides capable of binding to a specific peptide sequence in order to generate rationally designed anti-peptide peptides (Hurby et al., "Application of Synthetic Peptides: Antisense Peptides", In Synthetic Peptides, A User's Guide, W.H. Freeman, NY, pp. 289-307, 1992; Kaspaczak et al., Biochemistry 28:9230-9238, 1989).

Anti-peptide peptides can be generated by replacing the basic amino acid residues found in the peptide sequences of the kinases of the invention with acidic residues, while maintaining hydrophobic and uncharged polar groups. For example, lysine, arginine, and/or histidine residues are

replaced with aspartic acid or glutamic acid and glutamic acid residues are replaced by lysine, arginine or histidine.

The present invention also encompasses a method of detecting a STE20-related kinase polypeptide in a sample, comprising: (a) contacting the sample with an above-described antibody, under conditions such that immunocomplexes form, and (b) detecting the presence of said antibody bound to the polypeptide. In detail, the methods comprise incubating a test sample with one or more of the antibodies of the present invention and assaying whether the antibody binds to the test sample. Altered levels of a kinase of the invention in a sample as compared to normal levels may indicate disease.

Conditions for incubating an antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the antibody used in the assay. One skilled in the art will recognize that any one of the commonly available immunological assay formats (such as radioimmunoassays, enzyme-linked immunosorbent assays, diffusion based Ouchterlony, or rocket immunofluorescent assays) can readily be adapted to employ the antibodies of the present invention. Examples of such assays can be found in Chard ("An Introduction to Radioimmunoassay and Related Techniques" Elsevier Science Publishers, Amsterdam, The Netherlands, 1986), Bullock et al. ("Techniques in Immunocytochemistry," Academic Press, Orlando, FL Vol. 1, 1982; Vol. 2, 1983; Vol. 3, 1985), Tijssen ("Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands, 1985).

The immunological assay test samples of the present invention include cells, protein or membrane extracts of

cells, or biological fluids such as blood, serum, plasma, or urine. The test samples used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is testable with the system utilized.

A kit contains all the necessary reagents to carry out the previously described methods of detection. The kit may comprise: (i) a first container means containing an above-described antibody, and (ii) second container means containing a conjugate comprising a binding partner of the antibody and a label. In another preferred embodiment, the kit further comprises one or more other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound antibodies.

Examples of detection reagents include, but are not limited to, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the chromophoric, enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. The compartmentalized kit may be as described above for nucleic acid probe kits. One skilled in the art will readily recognize that the antibodies described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

VI. Isolation of Compounds Which Interact With STE20-Related Kinases

The present invention also relates to a method of detecting a compound capable of binding to a STE20-related kinase of the invention comprising incubating the compound with a kinase of the invention and detecting the presence of the compound bound to the kinase. The compound may be present within a complex mixture, for example, serum, body fluid, or cell extracts.

The present invention also relates to a method of detecting an agonist or antagonist of kinase activity or kinase binding partner activity comprising incubating cells that produce a kinase of the invention in the presence of a compound and detecting changes in the level of kinase activity or kinase binding partner activity. The compounds thus identified would produce a change in activity indicative of the presence of the compound. The compound may be present within a complex mixture, for example, serum, body fluid, or cell extracts. Once the compound is identified it can be isolated using techniques well known in the art.

The present invention also encompasses a method of agonizing (stimulating) or antagonizing kinase associated activity in a mammal comprising administering to said mammal an agonist or antagonist to a kinase of the invention in an amount sufficient to effect said agonism or antagonism. A method of treating diseases in a mammal with an agonist or antagonist of STE20-related kinase activity comprising administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize STE20-related kinase associated functions is also encompassed in the present application.

In an effort to discover novel treatments for diseases, biomedical researchers and chemists have designed, synthesized, and tested molecules that inhibit the function of protein kinases. Some small organic molecules form a class of compounds that modulate the function of protein kinases. Examples of molecules that have been reported to inhibit the function of protein kinases include, but are not limited to, bis monocyclic, bicyclic or heterocyclic aryl compounds (PCT WO 92/20642, published November 26, 1992 by Maguire et al.), vinylene-azaindole derivatives (PCT WO 94/14808, published July 7, 1994 by Ballinari et al.), 1-cyclopropyl-4-pyridyl-quinolones (U.S. Patent No. 5,330,992), styryl compounds (U.S. Patent No. 5,217,999), styryl-substituted pyridyl compounds (U.S. Patent No. 5,302,606), certain quinazoline derivatives (EP Application No. 0 566 266 A1), seleoindoles and selenides (PCT WO 94/03427, published February 17, 1994 by Denny et al.), tricyclic polyhydroxylic compounds (PCT WO 92/21660, published December 10, 1992 by Dow), and benzylphosphonic acid compounds (PCT WO 91/15495, published October 17, 1991 by Dow et al.).

Compounds that can traverse cell membranes and are resistant to acid hydrolysis are potentially advantageous as therapeutics as they can become highly bioavailable after being administered orally to patients. However, many of these protein kinase inhibitors only weakly inhibit the function of protein kinases. In addition, many inhibit a variety of protein kinases and will cause multiple side-effects as therapeutics for diseases.

Some indolinone compounds, however, form classes of acid resistant and membrane permeable organic molecules. WO 96/22976 (published August 1, 1996 by Ballinari et al.) describes hydrosoluble indolinone compounds that harbor

tetralin, naphthalene, quinoline, and indole substituents fused to the oxindole ring. These bicyclic substituents are in turn substituted with polar moieties including hydroxylated alkyl, phosphate, and ether moieties. U.S. Patent Application Serial Nos. 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang et al. (Lyon & Lyon Docket No. 221/187) and 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang et al. (Lyon & Lyon Docket No. 223/298) and International Patent Publication WO 96/22976, published August 1, 1996 by Ballinari et al., all of which are incorporated herein by reference in their entirety, including any drawings, describe indolinone chemical libraries of indolinone compounds harboring other bicyclic moieties as well as monocyclic moieties fused to the oxindole ring. Applications 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang et al. (Lyon & Lyon Docket No. 221/187), 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang et al. (Lyon & Lyon Docket No. 223/298), and WO 96/22976, published August 1, 1996 by Ballinari et al. teach methods of indolinone synthesis, methods of testing the biological activity of indolinone compounds in cells, and inhibition patterns of indolinone derivatives.

Other examples of substances capable of modulating kinase activity include, but are not limited to, tyrphostins, quinazolines, quinoxolines, and quinolines. The quinazolines, tyrphostins, quinolines, and quinoxolines referred to above include well known compounds such as those described in the literature. For example, representative

publications describing quinazolines include Barker et al., EPO Publication No. 0 520 722 A1; Jones et al., U.S. Patent No. 4,447,608; Kabbe et al., U.S. Patent No. 4,757,072; Kaul and Vougioukas, U.S. Patent No. 5, 316,553; Kreighbaum and Comer, U.S. Patent No. 4,343,940; Pegg and Wardleworth, EPO Publication No. 0 562 734 A1; Barker et al., Proc. of Am. Assoc. for Cancer Research 32:327 (1991); Bertino, J.R., Cancer Research 3:293-304 (1979); Bertino, J.R., Cancer Research 9(2 part 1):293-304 (1979); Curtin et al., Br. J. Cancer 53:361-368 (1986); Fernandes et al., Cancer Research 43:1117-1123 (1983); Ferris et al. J. Org. Chem. 44(2):173-178; Fry et al., Science 265:1093-1095 (1994); Jackman et al., Cancer Research 51:5579-5586 (1981); Jones et al. J. Med. Chem. 29(6):1114-1118; Lee and Skibo, Biochemistry 26(23):7355-7362 (1987); Lemus et al., J. Org. Chem. 54:3511-3518 (1989); Ley and Seng, Synthesis 1975:415-522 (1975); Maxwell et al., Magnetic Resonance in Medicine 17:189-196 (1991); Mini et al., Cancer Research 45:325-330 (1985); Phillips and Castle, J. Heterocyclic Chem. 17(19):1489-1596 (1980); Reece et al., Cancer Research 47(11):2996-2999 (1977); Sculier et al., Cancer Immunol. and Immunother. 23:A65 (1986); Sikora et al., Cancer Letters 23:289-295 (1984); Sikora et al., Analytical Biochem. 172:344-355 (1988); all of which are incorporated herein by reference in their entirety, including any drawings.

Quinoxaline is described in Kaul and Vougioukas, U.S. Patent No. 5,316,553, incorporated herein by reference in its entirety, including any drawings.

Quinolines are described in Dolle et al., J. Med. Chem. 37:2627-2629 (1994); MaGuire, J. Med. Chem. 37:2129-2131 (1994); Burke et al., J. Med. Chem. 36:425-432 (1993); and Burke et al. BioOrganic Med. Chem. Letters 2:1771-1774

(1992), all of which are incorporated by reference in their entirety, including any drawings.

Tyrphostins are described in Allen et al., Clin. Exp. Immunol. 91:141-156 (1993); Anafi et al., Blood 82:12:3524-3529 (1993); Baker et al., J. Cell Sci. 102:543-555 (1992); Bilder et al., Amer. Physiol. Soc. pp. 6363-6143:C721-C730 (1991); Brunton et al., Proceedings of Amer. Assoc. Cancer Resch. 33:558 (1992); Bryckaert et al., Experimental Cell Research 199:255-261 (1992); Dong et al., J. Leukocyte Biology 53:53-60 (1993); Dong et al., J. Immunol. 151(5):2717-2724 (1993); Gazit et al., J. Med. Chem. 32:2344-2352 (1989); Gazit et al., " J. Med. Chem. 36:3556-3564 (1993); Kaur et al., Anti-Cancer Drugs 5:213-222 (1994); Kaur et al., King et al., Biochem. J. 275:413-418 (1991); Kuo et al., Cancer Letters 74:197-202 (1993); Levitzki, A., The FASEB J. 6:3275-3282 (1992); Lyall et al., J. Biol. Chem. 264:14503-14509 (1989); Peterson et al., The Prostate 22:335-345 (1993); Pillemer et al., Int. J. Cancer 50:80-85 (1992); Posner et al., Molecular Pharmacology 45:673-683 (1993); Rendu et al., Biol. Pharmacology 44(5):881-888 (1992); Sauro and Thomas, Life Sciences 53:371-376 (1993); Sauro and Thomas, J. Pharm. and Experimental Therapeutics 267(3):119-1125 (1993); Wolbring et al., J. Biol. Chem. 269(36):22470-22472 (1994); and Yoneda et al., Cancer Research 51:4430-4435 (1991); all of which are incorporated herein by reference in their entirety, including any drawings.

Other compounds that could be used as modulators include oxindolinones such as those described in U.S. patent application Serial No. 08/702,232 filed August 23, 1996, incorporated herein by reference in its entirety, including any drawings.

VII. Biological Significance, Applications and Clinical Relevance of Novel STE20-Related Kinases

Human STLK2, STLK3, STLK4, STLK5, STLK6, and STLK7

STLK2, STLK4, STLK5, STLK6 and STLK7 belong to an expanding family of intracellular STKs that have varying degrees of sequence homology to SOK-1, a kinase implicated in oxidative stress agents (Pombo, CM et al, EMBO J. (17) 4537-4546, 1996). Our data shows that STLK2 is expressed highly in hematopoietic cells. Therefore, STLK2 may participate in the oxidative response pathway during inflammation. In addition, STLK2 could also be a possible component in the signaling pathways leading to T cell activation. High levels of STLK2 in several tumor cell lines could also imply that STLK2 might be involved in tumorigenesis.

STLK2 is most closely related to two human STE20-subfamily kinases: MST3 and SOK-1. MST3 is a 52,000 daltons cytoplasmic kinase that is ubiquitously expressed with its highest levels of expression found in heart, skeletal muscle and pancreas. The serine/threonine kinase activity of MST3 is activated by phosphorylation. Unlike SOK-1, MST3 prefers Mn⁺⁺ over Mg⁺⁺ and can use both GTP and ATP as phosphate donors. MST3 may undergo dimerization. No agonists have yet been identified that activate MST3. The downstream signaling mechanism of this kinase is unknown (Schinkmann, K and Blenis, J. (1997) J. Biol. Chem. 272, 28695-28703).

SOK-1 is a 50,000 daltons cytoplasmic kinase expressed predominantly in testis, large intestine, brain and stomach and to a lesser extent in heart and lung. SOK-1 is also expressed in the germinal center B-cell line (RAMOS) and in a mature B cell line (HS Sultan). The serine/threonine kinase activity of SOK-1 is activated by phosphorylation.

The C-terminus of SOK-1 has been shown to be inhibitory to the catalytic activity of this kinase. The only agonists known to activate SOK-1 are oxidant agents, like H₂O₂ and menadione, a quinone that is a potent intracellular generator of reactive oxygen species (Pombo, C.M. et al. EMBO J. 15, 4537-4546). SOK-1 is also activated by chemical anoxia through the generation of reactive oxygen species and release of calcium into the cytoplasm from intracellular stores. SOK-1, therefore, may play an important role in ischemia, the cause of myocardial infarction, stroke and acute renal failure (Pombo, C.M. et al. J. Biol. Chem. 272, 29372-29379 (1997)). The activity of SOK-1 in the response to oxidant stress is inversely correlated with the activity of the stress-activated protein kinases (SAPKs): elevated SOK-1 activity correlates with absent SAPK activity and vice-versa. SOK-1 does not activate any of the four MAP kinase pathways, SAPKs, p38, ERK-1 or MEK-5/ERK-5 (Pombo, C.M. et al. EMBO J. 15, 4537-4546). The downstream signaling mechanism of this kinase remains unknown.

STLK2 is expressed in a wide variety of immune cell types and tissues including thymus, dendrocytes, mast cells, monocytes, B cells (primary, Jurkat, RPMI, SR), T cells (CD8/CD4+, TH1, TH2, CEM, MOLT4) and megakaryocytes (K562), whereas STLK3 is restricted to thymus and STLK4 is predominately expressed in thymus, T cells (CD4/CD8+, TH1, CEM) and B cells (Jurkat, RPMI). Consequently, these STKs might participate in the oxidative response pathway during inflammation, reperfusion injury (stroke, surgery, shock), TNF α -mediated signaling, insulin desensitization, atherogenesis, vascular injury, T or B cell costimulation, or alternatively, participate in other MAPK-related signal transduction processes.

STLK5 is more distantly related to this STE20-subfamily including SOK-1 and STLK2, STLK3 and STLK4. STLK5, may therefore mediate a signaling pathway that is distinct from the oxidative stress response pathway.

5 The high degree of sequence homology in the C-termini of SOK-1, STLK2, STLK3, STLK4, STLK5, and STLK6 raises the possibility that these novel STKs, like SOK-1, may be subject to autoinhibition through a conserved C-terminal motif.

10

Human ZC1, ZC2, ZC3 and ZC4

ZC1 is a good candidate for any disease in which tyrosine kinase, cytokine, or heterotrimeric G-protein coupled receptors have been implicated. The mouse homologue binds to NCK, and is recruited to activated PDGF (Su et al., EMBO 16: 1279-1290, 1997). The Drosophila homolog has been shown to bind to TRAF2, implicating it in TNF- α signaling (Liu et al., (1999) Curr. Biol. 9:101-104, 1999)). While ZC1 does not contain the exact NCK- and TRAF2-binding domains, it is likely to bind to related proteins.

20 Of the ZC subfamily of STE20-related protein kinases, ZC1 has very broad over-expression in many tumor types, suggesting that it may be involved in cellular growth, transformation, or tumor progression. A truncated form of ZC1 containing only the C-terminal putative MEKK1-binding domain was found to reduce the number of foci generated by H-Ras-V12 in Rat Intestinal Epithelial cells (RIE-1). These data indicate that ZC1 may play a role in the ability for these cells to overcome contact inhibition and anchorage-dependent growth.

30 The ZC1 homolog, Misshapen (*msn*) in *Drosophila melanogaster* was cloned as a result of complementing a mutation in a developmental pathway required for dorsal

closure, a process involving changes in cell shape and position in the embryo (Treisman et al. *Gene* 186: 119-125, 1997). A *D. melanogaster* homolog of the JNK1/JNK2 kinases from mammals was shown to function downstream of *msn* in the dorsal-closure signaling pathway (Su et al. *Genes Dev.* 12:2371-2380, 1998).

While ZC1 could be involved in multiple aspects of tumorigenesis, by analogy with *Drosophila*, the role of misshapen in dorsal closure suggests a critical role in the regulation of the cytoskeleton for the processes of cell attachment, cell movement and perhaps migration.

The association of the ZC1 family members *msn* and NIK with TRAF2 may indicate a role for this kinase in cell survival and/or in apoptosis. The ZC1 family contains a highly conserved domain that in the mouse homolog, NIK, has been shown to bind to MEKK1 (Mitogen-activated/Extracellular-regulated Kinase Kinase 1) (Su et al., (1997) EMBO 16(6): 1279-90). MEKK1 is involved in cell survival and/or apoptosis in several systems (Schlesinger et al., *Front. Biosci.* 3:D1181-6, 1998). Depending on the context, MEKK1 appears to be upstream of either the ERK1/MAPK or the JNK/SAPK pathway [Schlesinger et al., (1998) *Front. Biosci.* 3:D1181-6]. Three homologues of ZC1: murine NIK (NCK-interacting kinase) (Su et al. EMBO 16:1279-90, 1997), *Drosophila msn* (Liu et al. *Curr. Biol.* 9:101-104, 1999) and human HGK (HPK/GCK-like kinase) (Yao et al., *J. Biol. Chem.* 274:2118-25, 1999) have all been shown to activate the JNK pathway when over-expressed in 293T cells.

ZC1 shares a high degree of homology with these other family members in both the kinase domain and the "MEKK"-binding domains, yet it differs in the intervening region, which contains several putative binding domains for upstream signaling adapter molecules (e.g. NCK, TRAF2). Unlike the

other family members, ZC1 does not appear to activate the JNK pathway in 293T cells as seen by its ability to induce expression of either a JUN or ATF2-driven luciferase gene. Upon co-transfection into these cells with HA-tagged JNK, modest activation of JNK was detected. ZC1 also modestly activated co-transfected ERK1. Both the ERK and the JNK activation were very slight compared with the positive controls in the assay (activated forms of MEK1 and MEKK1, respectively). In both cases, activation required the full-length kinase. While the kinase domain alone is up to 5x more active in autophosphorylation and in phosphorylation of MBP, it does not lead to activation of these potential downstream kinases. Based on the strong sequence homology of ZC1 with the other family members, it is very likely that ZC1 will be important for either JNK or ERK activation once the proper context is found.

ZC1 profoundly inhibits ERK1 kinase expression in co-transfection assays. This effect is dependent on ZC1 kinase activity, occurring with the wild-type and the kinase domain alone, but not with the kinase-dead mutant even though all three forms of ZC1 are expressed at similar levels. This may suggest a role for this kinase in transcriptional or post-transcriptional regulation.

ZC1 may be an important component in the signaling pathways mediated by the co-stimulatory receptor CD28 in T cells and/or by the pro-inflammatory cytokine TNF α , since co-transfection of the wild-type ZC1 activated the RE/AP-luciferase and NF κ B-luciferase reporter genes. While our data showed that ZC1 strongly activates NF κ B in T-cells, no activation of NF κ B driven luciferase was detectable in NIH 3T3 cells. A recent paper (J. Biol. Chem. 274:2118-25; 1999.) has shown that a human ZC1 splicing isoform, HGK, is involved in the TNF α -signaling pathways.

Given the importance of T cell activation in autoimmunity and transplantation, as well as the key role that TNF α plays in inflammatory diseases, it is possible that ZC1 could be a therapeutic target for immunological diseases which include but are not limited to: rheumatoid arthritis, chronic inflammatory bowel diseases (ie Crohn's disease), chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, and autoimmunity as well as organ transplantation and cardiovascular diseases.

ZC1 appears to be the human orthologue of murine NIK and possibly an orthologue of a *C. elegans* STE20-subfamily kinase encoded by the ZC504.4 cosmid.

Murine NIK is a 140,000 daltons kinase that is most highly expressed in brain and heart. NIK interacts with the SH3 domains of the adaptor molecule Nck through its proline-rich regions found in the C-terminal extra-catalytic region.

The specific regions that mediate this interaction are two PxxP motifs that are nearly uniformly conserved between NIK, ZC1,2,3 and the *C. elegans* STE20 ZC504.4 kinase. In addition, NIK binds MEKK1 through its 719 amino acid C-terminal (Su, Y-C. et al. (1997) EMBO J. 16, 1279-1290). MEKK1 is a membrane-associated kinase responsible for activating MKK4 (also known as SEK1), which in turn activates SAPK (Yan, M et al. (1994) Nature, 372, 798-800).

NIK may function as a kinase that links growth factor activated pathways and the stress-response pathway mediated by SAPKs. According to this hypothesis, activation of growth factor receptors leads to receptor tyrosine phosphorylation, Nck binding to the phosphorylated tyrosines via its SH2 domain, NIK redistribution to a membrane compartment via binding to the SH3 domain of Nck, and juxtaposition to the membrane-associated MEKK1. The NIK-

MEKK1 interaction would, in this fashion, turn on the SAPK pathway in response to growth factor stimulation (Su, Y-C. et al. (1997) EMBO J. 16, 1279-1290).

Given the high homology between ZC1, ZC2, ZC3, and ZC4 STKs and NIK, it is conceivable that these kinases may each function to connect growth factor- and stress-activated signaling pathways. The heterogeneity that the ZC kinases exhibit within their putative SH3-binding domain could provide signaling specificity in terms of the nature of the adaptor molecule that they bind. The high level of sequence conservation in the C-termini of the ZC1, ZC2 and ZC3 strongly suggests that these human kinases, like murine NIK, also may bind to MEKK1 and activate SAPKs. The ZC kinases also display strong homology at their C-termini to protein domains that bind small GTPase proteins such as Rab, Rho and Rac. For example, the C-termini of ZC1 is 36.2% identical to citron, a murine Rho-binding protein, and 23.1% identical to the rab-binding region of GC kinase. This suggests that, in addition to adaptor molecules, small GTPase proteins may also mediate membrane association and activation of the ZC kinases. The presence of a potential coiled-coil region located immediately C-terminal to the catalytic region strongly suggests that the ZC kinases may also be subject to regulation via homo or heterodimerization events.

The *C. elegans* STE20/ZC504.4 kinase is the product of the mig-15 gene. The product of this gene has been implicated in several developmental processes such as epidermal development, Q neuroblast migrations and muscle arm targeting in the developing worm (Zhu, X. and Hedgecock E. (1997) Worm Breeder's Gazette 14, 76). The high level of sequence conservation between the ZC kinases and the ZC504.4 *C. elegans* kinase will make *C. elegans* a valuable model

organism to study, through epistatic analysis, the signaling properties of the human ZC kinases.

Human KHS2

5 KHS1 (kinase homologous to SPS1/STE20) is a 100,000 dalton cytoplasmic STK that is expressed ubiquitously. KHS1 has been implicated in the mechanism of SAPK activation in response to inflammatory cytokines such as TNF α as well as to ultraviolet light, which also uses the TNF signaling pathway. TNF α binding to its receptors (TNFR1 and TNFR2) results in the sequential association with the receptor C-tail of multiple signaling molecules including TNFR1-associated death domain protein (TRADD), Fas-associated death domain protein (FADD or MORT1), TNFR-associated factor 2 (TRAF2), and the STK RIP (receptor interacting protein).
10 The TRADD-TRAF2 interaction is mediated by a conserved region present at the C-terminus of TRAF2, the TRAF domain.
15

Activation of the NF κ B and SAPK pathways is mediated by the ring finger motif present at the N-terminus of TRAF2 (Curr. Opinion in Cell. Biol. (1997) 9:247-251). KHS1 is activated by TNF α stimulation in a TRAF2-dependant manner and inhibition of KHS1 blocks TNF α -induced SAPK activation but not NF κ B activation. The mechanism by which TRAF2 activates KHS1 is not known. Cotransfection of TRAF2- and KHS1-expressing constructs in 293T cells failed to reveal a direct association between these two molecules. KHS1 activates the SAPK pathway by a direct association with the constitutively active kinase MEKK1. MEKK1 subsequently activates SEK1, which in turn activates SAPK. Neither the MAPK nor the p38 kinase pathways are activated by KHS1 (Shi, C-S and Kehrl. J.H. (1997) J. Biol. Chem. 272, 32102-32107). In addition to its catalytic domain, downstream

signaling of KHS1 requires its conserved C-terminus (Diener, K. et al (1997) Proc. Natl. Acad. Sci. 94, 9687-9692).

GCK (germinal center kinase) is a constitutively active 97,000 dalton STK that is broadly expressed. GCK may 5 participate in B-cell differentiation since its expression is localized to the germinal center within lymphoid follicles. GCK activates the SAPK pathway in response to TNF α via activation of SEK1. The upstream activators of GCK in response to cytokines as well as the immediate downstream 10 target of this kinase are unknown. The C-terminus of GCK is sufficient to activate SEK1 (Pombo, C.M. et al (1995) Nature, 377, 750-754).

The murine orthologue of GCK, rab8ip (rab8-interacting protein), is a 97,000 dalton protein that fractionates with 15 both the soluble cytoplasmic fraction as well as with a salt-sensitive fraction associated with the basolateral membrane of the trans-Golgi region in polarized MDCK epithelial cells. The C-terminus of rab8ip binds to rab8, a small GTP-binding protein required for vesicular transport 20 from the Golgi apparatus (Ren, M. et al. (1996) Proc. Natl. Acad. Sci. 93, 5151-5155). In addition to inducing the transcriptional activation of cytokines like IL2 via SAPK, GCK may also promote the rab-dependent release of secretory 25 proteins in response to TNF α (Buccione, R. et al (1995) Mol. Bio. Cell 6, 291).

HPK1 (hematopoietic protein kinase) is a constitutively active 90,000 dalton STK restricted to hematopoietic cells. HPK1 activates the SAPK pathway by 30 directly binding to and activating MEKK1 (Hu, M. et al (1996) Genes and Dev. 10:2251-2264) as well as the ubiquitously expressed mixed-lineage kinase MLK-3 (Kiefer, F. et al (1996) EMBO J. 15:7013-7025). This function of HPK1 requires, in contrast to GCK, both its kinase domain as

well as its C-terminus. The upstream activators of HPK1 remain unknown. HPK1 also plays a key role as a mediator of transforming growth factor- β (TGF β) signaling. HPK1 activates the TGF β -activated kinase (TAK), which in turn stimulates the SAPK pathway by phosphorylating SEK1 (Wang W. et al (1997) J. Biol. Chem. 272:22771-22775).

KHS2 is expressed in thymus, dendrocytes and monocytes. KHS2 could have a complementary function to that of KHS1 as a mediator of SAPK activation in the cellular response to inflammatory cytokines. KHS2 could have the potential to interact directly with TRAF2 since a STK with the predicted molecular weight of KHS2 (approximately 101,000 daltons) is found in the TNFR-TRAF2 complex upon TNF α stimulation (VanArsdale, T. and Ware, C.F. (1994) J. Immunol. 153, 3043-3050). The presence of a putative binding domain for Rab or a Rab-like molecule at the C-terminus of KHS2 indicates that KHS2, in addition to having a potential role in the TRAF2-dependant TNF α cytokine response, could also mediate signaling events that utilize small GTPase proteins. Alternatively, the binding of a small GTPase protein to the C-terminus of KHS2 may be required for its potential TRAF2-dependant signaling to a downstream kinase such as MEKK1.

Human GEK2, SULU1 and SULU3

A recent report (Y-W Qian et al., Science 282:1701-1704, 1998) described xPlkk1 as the activator of Plx1 (the Xenopus Polo kinase). In Xenopus oocytes, the STK Plkk1 can phosphorylate and activate Plx1 STK (the mammalian Polo kinase or PLK). A dominant-negative (kinase-dead) form of xPlkk1 prevents Plx1 activation and delays germinal vesicle breakdown. Yet another unidentified kinase is probably responsible for xPlkk1 activation during mitosis.

The homology through the entire length of the xPlkk1 protein with GEK2 suggests that GEK2 might represent the human homologue for xPlkk1. Based on this, GEK2 might be upstream of PLK in mammalian cells. In addition, based on the phage display screen results using the SULU1 coiled-coil2 domain as bait, SULU1 might also interact in vivo with GEK2 and therefore regulate GEK2 (and/or SLK through the coiled-coil domain) activation leading to PLK activation and mitosis.

If such a cascade of events is required for mitosis in mammalian cells, interruption of this signaling cascade at any point might block mitosis and could be beneficial for cancer treatment.

A recently cloned STE20-subfamily kinase, rat TAO1, is most likely the rodent orthologue of human SULU3 (Hutchinson, M. et al. J. Biol. Chem 273:28625-28632, 1998). TAO1 activates MEK3, 4 and 6 in vitro, while in transfected cells it associates and activates only MEK3, resulting in phosphorylation and activation of p38. These results implicate TAO1 (SULU3) in the regulation of the p38 containing stress-responsive MAP kinase pathway.

Human SULU1 is weakly expressed in hematopoietic sources whereas SULU3 is found in B-cells and TH1-restricted T cells. These mammalian SULU STKs display strong homology to the *C. elegans* SULU kinase. The role that this kinase plays in nematode development is unknown. The strong sequence homology between the catalytic domain of mammalian SULU kinases and other STE20-subfamily kinases such as SOK-1 (human STE20) and KHS2 suggests that the mammalian kinases may participate in the stress-response pathway. The potential coiled-coil domains found at the C-terminus of the SULU kinases may play a role in the regulation of this kinase.

Murine LOK (lymphocyte-oriented kinase) is a constitutively activated STK of approximately 130,000 daltons that is predominantly expressed in spleen, thymus and bone marrow (Kuramochi, S. et al (1997) J. Biol. Chem. 272: 22679-22684) as well as in meiotic testicular and primordial germ cells. The LOK1 gene is located in chromosome 11 of the mouse near the wr locus, a region that is associated with reproductive and neurological defects (Yanagisawa, M. et al (1996) Mol. Reprod. and Dev. 45:411-420). LOK does not activate any of the known MAPK pathways (ERK, JNK and p38) nor the NFkB pathway. The upstream signaling elements of LOK as well as the extracellular stimuli that utilize this kinase to elicit a biological response are also unknown (Kuramochi, S. et al (1997) J. Biol. Chem. 272: 22679-22684).

Human GEK2 is highly related to murine LOK, but based on sequence divergence in the non-catalytic domain, it appears to be a distinct member of this STE20-subfamily. GEK2 may signal through a pathway that remains to be defined. The presence of potential coiled-coil regions at the C-terminus of GEK2 could play a key role in regulating the functions of this kinase.

Human PAK4 and PAK5

The p21 activated protein kinases (PAK) are a closely related subgroup of the STE20 family of serine/threonine kinases. Extensive genetic and biochemical analysis of the budding yeast STE20 has shown the critical role this serine/threonine kinase plays at the juncture of several important intracellular pathways required to appropriately respond to extracellular signals. STE20 links the transcriptional response by mediating the activation of the appropriate downstream MAPK pathway as well as coupling

changes in cellular morphology via its control of the actin cytoskeleton.

A hallmark of the PAK subgroup is their small G protein-binding domain (PBD) that confers G protein-dependent activation upon this group of kinases. Via the PBD, PAKs bind to activated small G proteins resulting in the derepression of the PAK's intrinsic kinase activity.

Until recently, there were three known PAK kinases: PAK1, a 68 kD protein whose expression is restricted expression to brain, muscle, and spleen; PAK2 (PAK1, PAK65), a 62 kD protein whose expression is ubiquitous; and PAK3, a 65kD protein whose expression is restricted to the brain. Similar to STE20, the mammalian PAKs (1,2, and 3) have been shown to respond to extracellular signals (growth factors, mitogens, cytokines and a variety of cellular stresses) (Bagrodia, et al. (1995). J. Biol. Chem. 270: 22731-22737; Zhang, S., et al. (1995). J. Biol. Chem. 270: 23934-23936, Frost, J. et al. (1998) J. Biol. Chem. 273: 28191-28198; Galisteo, M. et al. (1996) J. Biol. Chem. 271: 20997-21000), and are linked to TCR activation (Yablonski, D., et al. (1998) EMBO J. 17: 5647-5657), and heterotrimeric G protein-coupled receptors (Knaus, U. et al. (1995) Science 269: 221-223).

The PAKs were originally identified as effectors for members of the Rho family of small G proteins (such as Rac and Cdc42), hence their name, p21-activated kinases (PAK) (Manser et al Nature 367:40-46). The recruitment of the PAKs to the appropriate intracellular location is critical to their function. Attempts to elucidate the role played by PAKs in intracellular signaling and morphological changes is complicated due to the complex interactions by which they can be recruited by such factors as activated small G

proteins (rac, cdc42), adaptors (nck) and exchange proteins (PIX, Cool).

The adaptor molecule, Nck, is constitutively bound via its SH3 domain to the proline-rich motif in the N-terminal portion of PAK1. Binding of the Nck-PAK complex to activated growth factor receptors in response to growth factor stimulation provides a mechanism to link growth factor-stimulated and stress-response pathways (Galisteo, M. et al. (1996) J. Biol. Chem. 271:20997-21000).

The PBD found at the N-terminus of PAK1 is responsible for its high-affinity interaction with the GTP-bound forms of Cdc42 and Rac (Burbelo, P. et al. (1995) J. Biol. Chem. 270:29071-29074). The exact mechanism through which the small GTPases activate PAKs may involve, in part, association of the kinase with activated growth factor receptors through guanine nucleotide exchange factors (GEFs). GEFs activate small GTPases by catalyzing the formation of their GTP-bound state, thereby promoting their association with, and activation of, PAKs. The known mammalian PAK kinases, as well as *Drosophila* and *C. elegans* PAKs, all conserve an N-terminal extracatalytic motif responsible for a high-affinity interaction with the GEF, PIX. The PAK-Cdc42 interaction and subsequent PAKs occurs as a PIX/PAK complex (Manser, E. et al. (1998) Molecular Cell, 1, 183-192).

PAK signaling stimulated by heterotrimeric G proteins is mediated through the interaction between a short conserved amino acid region located at the C-terminus of PAK1 with the G-protein β -subunit (Leeuw, T. et al. (1998) Nature, 391: 191-195).

A variety of studies have indicated that the human PAKs are involved in mediating the activation of stress-activated protein kinase pathways (JNK and to lesser extent p38).

PAKs are also potential mediators in the crosstalk between the pathways regulated by the Rho family of small G proteins and the signaling pathways directly downstream of Ras leading to the activation of the ERK pathway (Bagrodia, et al. (1995). J. Biol. Chem. 270: 22731-22737; Zhang, S., et al. (1995). J. Biol. Chem. 270: 23934-23936; Brown, J., et al. (1996) Curr Biol. 6:598-60596; Frost, J., et al. (1996). Mol. Cell. Biol. 16: 3707-3713).

PAK1 has been implicated in phosphorylating a regulatory site in MEK1 that is necessary for MEK1's ability to interact with Raf1 (Frost, et al. (1997) EMBO J. 16:6426-6438). PAK3 has been shown to phosphorylate Raf1 on a site that is important for Raf1 activity (King, A., et al. (1998). Nature 396: 180-183).

PAKs play an important role in controlling morphological changes in cell shape mediated by the actin cytoskeleton. Such morphological changes are required for cellular functions ranging from cell division and proliferation to cell motility and vesicle transport. PAK activity has been implicated in the localized assembly (leading edge) and disassembly (retracting edge) of focal adhesions necessary for cell motility (Frost J. et al (1998) J. Biol. Chem. 273:28191-28198).

PAK2 may have a role in the morphological changes induced during apoptosis (Membrane and morphological changes in apoptotic cells regulated by caspase-mediated activation of PAK2. (Rudel, T. (1997) Science. 276:1571-4)), and PAK1 may be important in preventing apoptosis (Faure S, et al. (1997) EMBO J. (1997) 16:5550-61). In addition to overcoming mitogen- and anchorage-independent growth, tumor cells need to escape the programmed cell death that accompanies deregulated cell growth. Thus, inhibition of PAKs may be effective in triggering apoptosis in tumors.

A direct requirement for PAKs in the transformation of mammalian cells has been shown for PAK1 and PAK2. Kinase-dead alleles of PAK1 block ras transformation of RAT1 and Schwann cells (Tang, Y., et al. (1997) Mol. Cell. Biol. 17, 4454-4464). Dominant-negative alleles of PAK2 have been shown to interfere with ras-mediated transformation of mammalian cells (Osada, S., (1997) FEBS Lett 404:227-233)

Mutations in PAK3 have been implicated in nonsyndromic X-linked mental retardation suggesting a role for PAK3 in cognitive function (Allen, K. et al. (1998) Nat. Genet. 20: 25-30). PAK1 has been implicated in neurite outgrowth in PC12 cells (Daniels, R. et al. (1998) EMBO J. 17: 754-764; Nikolic, M. et al. (1998) Nature 395:194-198).

Finally, PAK-like STKs may also play a role in AIDS pathogenesis since the myristoylated 27kD membrane-associated HIV Nef gene product directly interacts with and activates these kinases via cdc42 and Rac. The Nef-mediated activation of PAK-like STKs correlates with the induction of high viral titers and the development of AIDS in infected hosts (Cullen, B. R. (1996) Curr. Biol. 6:1557-1559).

Our results show that PAK4 is expressed in thymus, dendrocytes, mast cells, monocytes, as well as in T cells (TH2-restricted cells and MOLT4) and the B cell line RPMI. PAK5 is found in mast cells and in the T cell line MOLT4. These data suggest potential roles for PAK4 and PAK5 in the immune system.

PAK4 and PAK5 share with the known PAKs a potential cdc42-binding motif at their N-termini. Both PAK4 and PAK5 display sequence homology in their C-termini to a motif responsible for an interaction between PAK1 and the β -subunit of heterotrimeric G-proteins (amino acid residues 665-676 in PAK4, and amino acid residues 386-398 in PAK5). Consequently, PAK4, and possibly PAK5, could mediate

signaling events originating from growth factors as well as from ligands that stimulate G-protein-linked receptors.

PAK4 conserves a leucine (leu 44), that when mutated to a phenylalanine renders the kinase activity of human PAK1 constitutively active, bypassing its cdc42-binding requirement for activation (Brown J. et al (1996) Current Biol. 6:598-605). PAK5 contains an isoleucine at the equivalent position. Therefore, the mechanism by which cdc42 potentially activates human PAK1, PAK4, and possibly PAK5, may be very similar.

PAK4 and PAK5 however, lack the PIX-binding motif, and consequently cdc42-activating GEFs other than PIX (for example Dbl and Cool) must be responsible for the activation of these kinases. Alternatively, PAK4 and PAK5 may be activated by another GTPase, such as Rac1 which uses the Tiam1 GEF for its activation to the GTP-bound state.

PAK4 and PAK5 also lack the PxxP motif responsible for the Nck-PAK1 association. Between the PBD or cdc42-binding N-terminal motifs and the putative GEF-binding regions, PAK4 and PAK5 have long insertions (185 and 123 amino acids for PAK4 and PAK5, respectively) relative to PAK1. This region probably confers different binding characteristics to adaptor molecules and/or GEFs from those exhibited by known mammalian PAKs.

PAKs have been shown to be upstream in pathways leading to activation of both the JNK (Bagrodia, S., et al. (1995) J. Biol. Chem. 270: 22731-22737) and ERK kinase pathways (Brown, J., et al. (1996). Curr Biol. 6:598-605). PAK1 was shown to synergize with ras in activation of the ERK pathway through phosphorylation of MEK1 (Frost, J. et al. (1997). EMBO J. 16:6426-6438). Our data shows that MEK1 serves as an *in vitro* substrate for PAK4, suggesting a potential role

for PAK4 in the activation of the ERK pathway and mitogenesis.

PAK5 may also have a mitogenic role, and be linked to cancer, based on its expression profile (elevated RNA and protein levels in a wide variety of tumor cell lines), its interaction with cdc42 via its PBD, and the ability of a kinase-dead allele (Lys350, 351 Ala) to block ras transformation of NIH3T3 cells. Thus, a screen for small molecule inhibitors of PAK5 kinase activity may yield compounds with therapeutic potential for intervention in cancer derived from a wide variety of tissue types.

PAK4 and PAK5 may also play a role in HIV pathogenesis as potential mediators of Nef signaling, since none of the known PAKs correspond to the PAK-like kinase shown to interact with, and be activated by, the HIV nef protein (Lu, X. et al. (1996) Current Biology 6:1677-1684)

The 3' untranslated region of PAK4 contains a CA repeat that is prone to undergo expansion. CA dinucleotide repeat instability has been associated with disease (Toren, M.Z. et al (1998) Am. J. Hematol. 57: 148-152), and expansion of such repeat in the 3' untranslated region of PAK4 could implicate this kinase in as yet unknown pathologies.

Clinical applications

Human STLK2, STLK3, STLK4, STLK5, STLK6, and STLK7

STLK3, STLK5, STLK6 and STLK7, as well as other homologues of the STLK subfamily of STE20 protein kinases such as STLK4, may play an important role as mediators of the immune response. Thus, they are targets for the development of specific small molecule inhibitors to treat immunological diseases, including, but not limited to, rheumatoid arthritis, chronic inflammatory bowel diseases (e.g. Crohn's disease), chronic inflammatory pelvic disease,

multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis and autoimmunity, as well as in organ transplantation. Other diseases include cardiovascular diseases.

5 The human STLKs may also play an important role in cell growth regulation. Thus, they are targets for developing small molecule kinase inhibitors for the treatment of cancer and metastases. STLK5 maps to a chromosomal region frequently amplified in a variety of tumors including those from non-small cell lung cancer, breast cancer and peripheral nerve tumors. This suggests that STLK5 could play a role in the development, maintenance, or progression of human tumors.

10

15 The potential role of human STLKs 2, 3, and 4 in mediating oxidative stress strongly suggests that drugs targeting these kinases could prove useful in the treatment of myocardial infarction, arrhythmia and other cardiomyopathies, stroke, renal failure, oxidative stress-related neurodegenerative disorders such amyotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome, a necrotizing mitochondrial encephalopathy, as well.

20

Human ZC1, ZC2, ZC3, and ZC4

25 ZC1 may be a component of the CD28-signaling pathway and therefore important in T cell activation. As such, ZC1 as well as other ZC subfamily kinases, are targets for the development of specific small molecule inhibitors to treat immunological diseases, including, but not limited to, rheumatoid arthritis, chronic inflammatory bowel diseases (e.g. Crohn's disease), chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis and autoimmunity, as well as organ

30

transplantation. Other diseases include cardiovascular diseases.

ZC1 and ZC2 are also implicated in cell growth regulation. Thus, ZC subfamily kinases are targets for developing small molecule inhibitors for the treatment of cancer and metastases. ZC2 maps to a chromosomal region frequently amplified in a variety of tumors including those from non-small cell lung cancer, small cell lung cancer, and cervical cancer. This suggests that ZC2 could play a role in the development, maintenance, or progression of human tumors.

The role of human ZC1, ZC2, ZC3, and ZC4 in the inflammatory and stress-response pathways, strongly suggests that drugs targeting these kinases could have strong immunosuppressive actions. These drugs can prove valuable for the treatment of rheumatoid arthritis, atherosclerosis, autoimmune disorders and organ transplantation among others.

At least one very important class of immunosuppressants, corticosteroids, functions by blocking SAPK activation at an as yet undefined site on this pathway (Swantek, J.L. et al (1997) Mol. Cell. Biol. (1997) 6274-6282). Other immunosuppressive drugs like the pyridinyl imidazoles specifically target the p38 kinases (Kumar, S. et al (1997) Biochem. Biophys. Res. Commun. 235: 533-528). Drug targeting of the MAPK and p38 pathways could lead to the development of novel immunosuppressants.

Human SULU and GEK

The potential role of these novel STE20-related protein kinases in the control of mitosis strongly suggests that agents that specifically inhibit these kinases could be useful for cancer and metastases treatment.

The close homology of human STLK5, GEK2, SULU1 and SULU3 to STE20-subfamily kinases involved in the stress and oxidative response pathway strongly suggests that drugs targeting these kinases may also be useful as immunosuppressants as well as to treat ischemic disorders.

Human KHS2

The role of human KHS2 in the inflammatory and stress-response pathways, strongly suggests that drugs targeting this and related kinases could have strong immunosuppressive actions. These drugs can prove valuable for the treatment of rheumatoid arthritis, atherosclerosis, autoimmune disorders and organ transplantation among others. At least one very important class of immunosuppressants, corticosteroids, functions by blocking SAPK activation at an as yet undefined site on this pathway (Swantek, J.L. et al (1997) Mol. Cell. Biol. (1997) 6274-6282). Other immunosuppressive drugs like the pyridinyl imidazoles specifically target the p38 kinases (Kumar, S. et al (1997) Biochem. Biophys. Res. Commun. 235: 533-528). Drug targeting of the MAPK and p38 pathways could lead to the development of novel immunosuppressants.

Human PAK family

PAK5 has a role in cancer based on its expression profile (elevated RNA and protein levels in wide variety of tumor lines), its interaction with Cdc42 via its PBD, and the ability of the kinase-dead allele of PAK5 (Lys350, 351Ala) to block ras transformation of NIH3T3 cells. Thus, a screen for small molecule inhibitors of PAK5 kinase activity may yield compounds with therapeutic potential for intervention in cancers and metastases derived from a wide range of tissue types.

PAK5 maps to a chromosomal region frequently amplified in a variety of tumors including those from non-small cell lung cancer, and small cell lung cancer. These findings suggest that PAK5 could play a role in the development, maintenance, or progression of human tumors and/or metastases.

The role of human PAK4, and PAK5 in the inflammatory and stress-response pathways also strongly suggests that drugs targeting these kinases could have strong immunosuppressive actions. These drugs can prove valuable for the treatment of rheumatoid arthritis, atherosclerosis, autoimmune disorders and organ transplantation among others.

At least one very important class of immunosuppressants, corticosteroids, functions by blocking SAPK activation at an as yet undefined site on this pathway (Swantek, J.L. et al (1997) Mol. Cell. Biol. (1997) 6274-6282). Other immunosuppressive drugs like the pyridinyl imidazoles specifically target the p38 kinases (Kumar, S. et al (1997) Biochem. Biophys. Res. Commun. 235: 533-528). Drug targeting of the MAPK and p38 pathways could lead to the development of novel immunosuppressants. In addition, drugs targeting PAK4 or PAK5 could prove useful as immunosuppressants as well as in AIDS treatment.

25 VIII. Transgenic Animals.

A variety of methods are available for the production of transgenic animals associated with this invention. DNA can be injected into the pronucleus of a fertilized egg before fusion of the male and female pronuclei, or injected into the nucleus of an embryonic cell (e.g., the nucleus of a two-cell embryo) following the initiation of cell division (Brinster et al., Proc. Nat. Acad. Sci. USA 82: 4438-4442, 1985). Embryos can be infected with viruses, especially

retroviruses, modified to carry inorganic-ion receptor nucleotide sequences of the invention.

Pluripotent stem cells derived from the inner cell mass of the embryo and stabilized in culture can be manipulated in culture to incorporate nucleotide sequences of the invention. A transgenic animal can be produced from such cells through implantation into a blastocyst that is implanted into a foster mother and allowed to come to term.

Animals suitable for transgenic experiments can be obtained from standard commercial sources such as Charles River (Wilmington, MA), Taconic (Germantown, NY), Harlan Sprague Dawley (Indianapolis, IN), etc.

The procedures for manipulation of the rodent embryo and for microinjection of DNA into the pronucleus of the zygote are well known to those of ordinary skill in the art (Hogan et al., *supra*). Microinjection procedures for fish, amphibian eggs and birds are detailed in Houdebine and Chourrout (*Experientia* 47: 897-905, 1991). Other procedures for introduction of DNA into tissues of animals are described in U.S. Patent No., 4,945,050 (Sandford et al., July 30, 1990).

By way of example only, to prepare a transgenic mouse, female mice are induced to superovulate. Females are placed with males, and the mated females are sacrificed by CO₂ asphyxiation or cervical dislocation and embryos are recovered from excised oviducts. Surrounding cumulus cells are removed. Pronuclear embryos are then washed and stored until the time of injection. Randomly cycling adult female mice are paired with vasectomized males. Recipient females are mated at the same time as donor females. Embryos then are transferred surgically. The procedure for generating transgenic rats is similar to that of mice (Hammer et al., *Cell* 63:1099-1112, 1990).

Methods for the culturing of embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation, calcium phosphate/DNA precipitation and direct injection also are well known to those of ordinary skill in the art (Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, E.J. Robertson, ed., IRL Press, 1987).

In cases involving random gene integration, a clone containing the sequence(s) of the invention is co-transfected with a gene encoding resistance. Alternatively, the gene encoding neomycin resistance is physically linked to the sequence(s) of the invention. Transfection and isolation of desired clones are carried out by any one of several methods well known to those of ordinary skill in the art (E.J. Robertson, *supra*).

DNA molecules introduced into ES cells can also be integrated into the chromosome through the process of homologous recombination (Capecchi, *Science* 244: 1288-1292, 1989). Methods for positive selection of the recombination event (*i.e.*, neo resistance) and dual positive-negative selection (*i.e.*, neo resistance and gancyclovir resistance) and the subsequent identification of the desired clones by PCR have been described by Capecchi, *supra* and Joyner et al. (*Nature* 338: 153-156, 1989), the teachings of which are incorporated herein in their entirety including any drawings. The final phase of the procedure is to inject targeted ES cells into blastocysts and to transfer the blastocysts into pseudopregnant females. The resulting chimeric animals are bred and the offspring are analyzed by Southern blotting to identify individuals that carry the transgene. Procedures for the production of non-rodent mammals and other animals have been discussed by others (Houdebine and Chourrout, *supra*; Pursel et al., *Science*

244:1281-1288, 1989; and Simms et al., Bio/Technology 6:179-183, 1988).

5

10

Thus, the invention provides transgenic, nonhuman mammals containing a transgene encoding a kinase of the invention or a gene effecting the expression of the kinase. Such transgenic nonhuman mammals are particularly useful as an *in vivo* test system for studying the effects of introduction of a kinase, or regulating the expression of a kinase (i.e., through the introduction of additional genes, antisense nucleic acids, or ribozymes).

15

20

A "transgenic animal" is an animal having cells that contain DNA which has been artificially inserted into a cell, which DNA becomes part of the genome of the animal which develops from that cell. Preferred transgenic animals are primates, mice, rats, cows, pigs, horses, goats, sheep, dogs and cats. The transgenic DNA may encode human STE20-related kinases. Native expression in an animal may be reduced by providing an amount of anti-sense RNA or DNA effective to reduce expression of the receptor.

25

IX. Gene Therapy

STE20-related kinases or their genetic sequences will also be useful in gene therapy (reviewed in Miller, Nature 357:455-460, 1992). Miller states that advances have resulted in practical approaches to human gene therapy that have demonstrated positive initial results. The basic science of gene therapy is described in Mulligan (Science 260:926-931, 1993).

30

In one preferred embodiment, an expression vector containing STE20-related kinase coding sequence is inserted into cells, the cells are grown *in vitro* and then infused in large numbers into patients. In another preferred embodiment, a DNA segment containing a promoter of choice (for

example a strong promoter) is transferred into cells containing an endogenous gene encoding kinases of the invention in such a manner that the promoter segment enhances expression of the endogenous kinase gene (for example, the promoter segment is transferred to the cell such that it becomes directly linked to the endogenous kinase gene).

The gene therapy may involve the use of an adenovirus containing kinase cDNA targeted to a tumor, systemic kinase increase by implantation of engineered cells, injection with kinase-encoding virus, or injection of naked kinase DNA into appropriate tissues.

Target cell populations may be modified by introducing altered forms of one or more components of the protein complexes in order to modulate the activity of such complexes. For example, by reducing or inhibiting a complex component activity within target cells, an abnormal signal transduction event(s) leading to a condition may be decreased, inhibited, or reversed. Deletion or missense mutants of a component, that retain the ability to interact with other components of the protein complexes but cannot function in signal transduction may be used to inhibit an abnormal, deleterious signal transduction event.

Expression vectors derived from viruses such as retroviruses, vaccinia virus, adenovirus, adeno-associated virus, herpes viruses, several RNA viruses, or bovine papilloma virus, may be used for delivery of nucleotide sequences (e.g., cDNA) encoding recombinant kinase of the invention protein into the targeted cell population (e.g., tumor cells). Methods which are well known to those skilled in the art can be used to construct recombinant viral vectors containing coding sequences (Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor

Laboratory, N.Y., 1989; Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, N.Y., 1989). Alternatively, recombinant nucleic acid molecules encoding protein sequences can be used as naked DNA or in a reconstituted system e.g., liposomes or other lipid systems for delivery to target cells (e.g., Felgner et al., Nature 337:387-8, 1989). Several other methods for the direct transfer of plasmid DNA into cells exist for use in human gene therapy and involve targeting the DNA to receptors on cells by complexing the plasmid DNA to proteins (Miller, *supra*).

In its simplest form, gene transfer can be performed by simply injecting minute amounts of DNA into the nucleus of a cell, through a process of microinjection (Capecci, Cell 22:479-88, 1980). Once recombinant genes are introduced into a cell, they can be recognized by the cell's normal mechanisms for transcription and translation, and a gene product will be expressed. Other methods have also been attempted for introducing DNA into larger numbers of cells.

These methods include: transfection, wherein DNA is precipitated with CaPO₄ and taken into cells by pinocytosis (Chen et al., Mol. Cell Biol. 7:2745-52, 1987); electroporation, wherein cells are exposed to large voltage pulses to introduce holes into the membrane (Chu et al., Nucleic Acids Res. 15:1311-26, 1987); lipofection/liposome fusion, wherein DNA is packaged into lipophilic vesicles which fuse with a target cell (Felgner et al., Proc. Natl. Acad. Sci. USA. 84:7413-7417, 1987); and particle bombardment using DNA bound to small projectiles (Yang et al., Proc. Natl. Acad. Sci. 87:9568-9572, 1990). Another method for introducing DNA into cells is to couple the DNA to chemically modified proteins.

It has also been shown that adenovirus proteins are capable of destabilizing endosomes and enhancing the uptake of DNA into cells. The admixture of adenovirus to solutions containing DNA complexes, or the binding of DNA to polylysine covalently attached to adenovirus using protein crosslinking agents substantially improves the uptake and expression of the recombinant gene (Curiel et al., Am. J. Respir. Cell. Mol. Biol., 6:247-52, 1992).

As used herein "gene transfer" means the process of introducing a foreign nucleic acid molecule into a cell. Gene transfer is commonly performed to enable the expression of a particular product encoded by the gene. The product may include a protein, polypeptide, anti-sense DNA or RNA, or enzymatically active RNA. Gene transfer can be performed in cultured cells or by direct administration into animals.

Generally gene transfer involves the process of nucleic acid contact with a target cell by non-specific or receptor mediated interactions, uptake of nucleic acid into the cell through the membrane or by endocytosis, and release of nucleic acid into the cytoplasm from the plasma membrane or endosome. Expression may require, in addition, movement of the nucleic acid into the nucleus of the cell and binding to appropriate nuclear factors for transcription.

As used herein "gene therapy" is a form of gene transfer and is included within the definition of gene transfer as used herein and specifically refers to gene transfer to express a therapeutic product from a cell *in vivo* or *in vitro*. Gene transfer can be performed *ex vivo* on cells which are then transplanted into a patient, or can be performed by direct administration of the nucleic acid or nucleic acid-protein complex into the patient.

In another preferred embodiment, a vector having nucleic acid sequences encoding a STE20-related kinase

polypeptide is provided in which the nucleic acid sequence is expressed only in specific tissue. Methods of achieving tissue-specific gene expression are set forth in International Publication No. WO 93/09236, filed November 3, 5 1992 and published May 13, 1993.

In all of the preceding vectors set forth above, a further aspect of the invention is that the nucleic acid sequence contained in the vector may include additions, deletions or modifications to some or all of the sequence of 10 the nucleic acid, as defined above.

In another preferred embodiment, a method of gene replacement is set forth. "Gene replacement" as used herein means supplying a nucleic acid sequence which is capable of being expressed *in vivo* in an animal and thereby providing 15 or augmenting the function of an endogenous gene which is missing or defective in the animal.

X. Administration of Substances

Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in U.S. Application Serial No. 08/702,282, filed August 23, 1996 and International patent publication number WO 96/22976, published August 1 1996, both of which are incorporated 20 herein by reference in their entirety, including any drawings, figures, or tables. Those skilled in the art will appreciate that such descriptions are applicable to the 25 present invention and can be easily adapted to it.

The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used, and the size and physiological condition of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from 30

cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the IC₅₀ as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors, and major organs can be also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection methods such as X-ray, CAT scan, and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharmacokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

Toxicity studies can also be carried out by measuring the blood cell composition. For example, toxicity studies can be carried out in a suitable animal model as follows: 1) the compound is administered to mice (an untreated control mouse should also be used); 2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and 3) the samples are analyzed for red and white blood cell counts, blood cell composition, and the percent of lymphocytes versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary

Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia, *Journal of American Veterinary Medical Assoc.*, 202:229-249, 1993).

5 Representative animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness, or toxicity. Gross abnormalities in tissue are noted, and tissues are examined histologically. Compounds causing a reduction in body weight or blood components are less preferred, as are
10 compounds having an adverse effect on major organs. In general, the greater the adverse effect the less preferred the compound.

For the treatment of cancers the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500
15 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness.

Plasma levels should reflect the potency of the drug.
20 Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

EXAMPLES

The examples below are not limiting and are merely
25 representative of various aspects and features of the present invention. The examples below demonstrate the isolation and characterization of the STE20-related kinases of the invention.

EXAMPLE 1: Isolation of cDNAs Encoding Mammalian
STE20-related Protein Kinases

Materials and Methods

Identification of novel clones

5 Total RNAs were isolated using the Guanidine Salts/Phenol extraction protocol of Chomczynski and Sacchi (P. Chomczynski and N. Sacchi, Anal. Biochem. 162, 156 (1987)) from primary human tumors, normal and tumor cell lines, normal human tissues, and sorted human hematopoietic 10 cells. These RNAs were used to generate single-stranded cDNA using the Superscript Preamplification System (GIBCO BRL, Gaithersburg, MD; Gerard, GF et al. (1989), FOCUS 11, 66) under conditions recommended by the manufacturer. A typical reaction used 10 µg total RNA with 1.5 µg oligo(dT)₁₂-₁₈ in a reaction volume of 60 µL. The product was treated with RNaseH and diluted to 100 µL with H₂O. For subsequent PCR amplification, 1-4 µL of this ssDNA was used in each reaction.

20 Degenerate oligonucleotides were synthesized on an Applied Biosystems 3948 DNA synthesizer using established phosphoramidite chemistry, precipitated with ethanol and used unpurified for PCR. The sequence of some of the degenerate oligonucleotide primers and the amino acid motif they encode is as follows:

25 TRK1 5'-CTGAATTGGNGCNTTYGGNAARGT-3' GAFGKV (sense)
 TRK4 5'-GCTGGATCCYTCNGGNNGCATCCA-3' WMPPE (antisense)
 ROS1 5'-GCNTTYGGNGARGTNTAYGARGG-3' AFGEVYEG (sense)
 CCK4b 5'-GCTGGATCCYTCNGGNSWCATCCA-3' WMSPE (antisense)
 CCK4c 5'-GAGTTYGGNGARGTNTYYTNGC-3' EFGEVYEG (sense)

30 These primers were derived from the sense and antisense strands of conserved motifs within the catalytic domain of several protein kinases. Degenerate nucleotide residue

designations are: N = A, C, G, or T; R = A or G; Y = C or T; H = A, C or T not G; D = A, G or T not C; S = C or G; and W = A or T.

PCR reactions were performed using degenerate primers applied to multiple single-stranded cDNAs. The primers were added at a final concentration of 5 μ M each to a mixture containing 10 mM TrisHCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 200 μ M each deoxynucleoside triphosphate, 0.001% gelatin, 1.5 U AmpliTaq DNA Polymerase (Perkin-Elmer/Cetus), and 1-4 μ L cDNA. Following 3 min denaturation at 95 °C, the cycling conditions were 94 °C for 30 s, 50 °C for 1 min, and 72 °C for 1 min 45 s for 35 cycles. PCR fragments migrating between 300-350 bp were isolated from 2% agarose gels using the GeneClean Kit (Bio101), and T-A cloned into the pCRII vector (Invitrogen Corp. U.S.A.) according to the manufacturer's protocol.

Colonies were selected for mini plasmid DNA-preparations using Qiagen columns and the plasmid DNA was sequenced using a cycle sequencing dye-terminator kit with AmpliTaq DNA Polymerase, FS (ABI, Foster City, CA). Sequencing reaction products were run on an ABI Prism 377 DNA Sequencer, and analyzed using the BLAST alignment algorithm (Altschul, S.F. et al., J.Mol.Biol. 215: 403-10).

Additional PCR strategies were employed to connect various PCR fragments or ESTs using exact or near exact oligonucleotide primers as detailed in the results section for each cDNA. PCR conditions were as described above except the annealing temperatures were calculated for each oligo pair using the formula: $T_m = 4(G+C) + 2(A+T)$.

30

Isolation of cDNA clones:

Human cDNA libraries were probed with PCR or EST fragments corresponding to STE20-related genes. Probes were

³²P-labeled by random priming and used at 2x10⁶ cpm/mL following standard techniques for library screening. Pre-hybridization (3 h) and hybridization (overnight) were conducted at 42 °C in 5X SSC, 5X Denhart's solution, 2.5% dextran sulfate, 50 mM Na₂PO₄/NaHPO₄, pH 7.0, 50% formamide with 100 mg/mL denatured salmon sperm DNA. Stringent washes were performed at 65 °C in 0.1X SSC and 0.1% SDS. DNA sequencing was carried out on both strands using a cycle sequencing dye-terminator kit with AmpliTaq DNA Polymerase, FS (ABI, Foster City, CA). Sequencing reaction products were run on an ABI Prism 377 DNA Sequencer.

Makegene Bioinformatics EST assembler

The EST reports were downloaded from ncbi (www.ncbi.nlm.nih.gov). After uncompressed the files, the program 'report2est' was scripted to extract the following information: 1) EST names, 2) GenBank Accession numbers, 3) GenBank gi numbers, 4) Clone Id numbers, 5) the nucleotide sequences of the ESTs 6) the organism, 7) the library name, 8) the name of the lab, and 9) the institution. The output of 'report2est' is a file in FASTA format with all of the information listed above in the first line of each entry except the sequence, which is listed in the second line of each entry. The resulting file is formatted for BLAST using 'pressdb' (available as part of the ncbi tool kit).

To build a gene or part of a gene from ESTs, the program 'makegene' was developed. Input to this program is a query sequence and the organism/species for which a gene is to be built. An initial search of the formatted EST database described above is performed using BLAST (blastn). Any results that contain warnings, such as polyA tails or other repeat elements, are eliminated from future queries. The program 'blast_parse_reports' was developed to extract

the FASTA header line from the search results and the output is then filtered to extract only FASTA header lines for the desired species.

The initial results, having been filtered for warnings and species, go into a loop in which searches against the database are repeated until no new ESTs are found. The loop consists of the following steps: 1) when possible the names of both ends of the ESTs are extracted from the database by searching using the 'Clone Id' field or the part of the 'EST name' field before the .r or .s postscript, 2) any ESTs that have been used as queries in previous loops are removed from the current query by the program 'subtract', 3) the resulting list of ESTs is used to extract the sequences from the database by the program batch_parse_fasta, 4) BLAST is run against the database using each sequence, 5) the output files from BLAST containing warnings are removed, 6) the results are filtered by species, and 7) the loop is reentered if there were new ESTs found in the previous pass through the loop.

The ESTs chosen by 'makegene' are used as input for the program 'mpd2_cluster' (Hide, W., Burke, J., and Davison, D. U. of Houston, unpublished) which clusters overlapping sequences. The programs 'contig' (Kerlavage, T., TIGR, unpublished), 'gde2mult' and 'gde2sing' (Smith, S.W., et al., CABIOS 10, 671-675 (1994)), are used to make an alignment and consensus sequence of the overlapping ESTs.

RESULTS

cDNA cloning and characterization of STLK2

The human STLK2 cDNA sequence is composed of two overlapping EST fragments, AA191319 and WI6504, that were identified using a Smith-Waterman search of the EST database with STLK1 (MST3 GB:AF024636) as a query. The complete

sequence of both clones was determined and used to generate the full-length human STL2 sequence.

EST clone AA191319 contains a 1327 bp insert and an ORF of 1146 bp (382 amino acids). EST clone W16504 contains a 2474 bp insert (not including the poly-A tail) and an ORF of 687 bp (382 amino acids).

The full-length human STLK2 cDNA (SEQ ID NO. 1) is 3268 bp long. AA191319 spans positions 1-1327 and W16504 positions 743-3216. The overlap between these two clones exhibits 100% sequence identity. The human STLK2 cDNA contains a 1248 bp ORF flanked by a 181 bp 5' UTR (1-181) and a 1784 bp 3' UTR (1433-3216) that is followed by a 52 nucleotide polyadenylated region. A polyadenylation signal (AATAAA) is found at positions 3193-3198. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for STLK2. Furthermore, human STLK2, and the related SOK-1 and MST3 proteins, conserve the amino acid sequence immediately following this presumed initiating methionine.

Several EST fragments span the complete STLK2 sequence with AA191319 at the 5' end and W16504 at the 3' end.

All searches against the public nucleic acid database (NRN) and protein database (NRP) were conducted using the Smith-Waterman gap alignment program ((Smith, TF and Waterman, MS (1981) J. Mol. Biol., 147, 195-197).) with the PAM100 matrix and gap open and extension penalties of 14:1, respectively.

30 cDNA Cloning and Characterization of STLK3

A mammalian STLK3 clone, 135-31-19, was first identified from a PCR screen with the degenerate oligos, TRK1 and TRK4, applied to a sscDNA generated from adult rat

brain substantia nigra. Sequence analysis of the 457 bp insert indicated that it represented a novel member of the STE20-subfamily of STKs.

5 A Smith-Waterman search of the EST database with the rat STLK3 fragment and human STLK1 (MST3 GB:AF024636) as queries identified several overlapping ESTs spanning most of the human STLK3 cDNA sequence. A Makegene analysis generated a 3037 bp contig from approximately 44 EST sequences. Since the 3' ESTs were not commercially available, a pair of primers (5'-CACAGAACGGTCAGATTCAC-3' and 5'-GATCAGGGTGACATCAAGGGAC-3') were derived from this region to generate PCR clone 3R21-20-6 from human fetal liver sscDNA. This clone and EST AA278967 were fully sequenced to generate the full-length STLK2 cDNA sequence.

10 15 AA278967 is a 837 bp EST isolated by the IMAGE consortium from cDNA made from CD20+/IgD- germinal center B cells sorted from human tonsillar cells.

20 PCR clone 3R21-20-6 was isolated from human fetal sscDNA and contains a 1116 bp insert, including a 1086 bp ORF encoding the 362 C-terminal amino acids of STLK3.

25 The full-length human STLK3 cDNA (SEQ ID NO. 2) is 3030 bp long. AA278967 spans positions 1-814 and 3R21-20-6 spans positions 464-1579. The overlap between these two clones exhibits 100% sequence identity. The remaining 1452 bp of 3' UTR is derived from an assembly of multiple unconfirmed EST fragments.

30 The near full-length human STLK3 cDNA (SEQ ID NO. 2) is 3030 bp long and consists of a 1548 bp ORF flanked by a 1476 bp 3' UTR (1550-3025) and a 5 nucleotide polyadenylated region. A polyadenylation signal (AATAAA) begins at position 3004. Since the coding region is open throughout the 5' extent of this sequence, this is apparently a partial cDNA clone lacking the N-terminal start methionine. Six

copies of a "GGCCCC" repeat were observed in positions 21-67. Five independent ESTs (AA150838, AA286879, AA251679, AA252004, AA278967) showed the same repeat, suggesting that this sequence may be an integral region of the human STLK3 gene. Stronger evidence for this being the case is provided by the sequence of the murine orthologue of STLK3 represented by a 876 bp EST W20737.

Multiple EST fragments span the complete STLK3 sequence with AA278967 at the 5' end and AA628477 and others at the 10 3' end.

cDNA Cloning and Characterization of STLK4

The human STLK4 cDNA sequence is composed of two overlapping EST fragments, AA297759 and AA100484, that were 15 identified using a Smith-Waterman search of the EST database with STLK1 (MST3 GB:AF024636) as a query. The complete sequence of both clones was determined and used to generate the near full-length human STLK4 sequence.

AA100484 is an IMAGE consortium cDNA clone isolated 20 from the T-84 colonic epithelium cell line. It has an insert of 3694 bp and a coding region of 1146 bp (382 amino acids). A Smith-Waterman sequence alignment against the NRN database showed this EST to be 71.4% identical to the human STE20-like kinase (GB:X99325).

W16504 is an IMAGE consortium clone isolated from a 25 human fetal heart cDNA library. It has an insert length of 2474 bp (not including the poly-A tail) and a coding region of 687 bp (229 amino acids). A Smith-Waterman sequence alignment of W16504 against the NRN database showed this EST 30 to be 69.2% identical to the human STE20-like kinase (GB:X99325).

The full-length human STLK2 cDNA (SEQ ID NO. 1) is 3268 bp long. AA191319 spans positions 1-1327, and W16504

positions 743-3216. The overlap between these two clones is 585 bp long with 100% sequence identity.

AA100484 is an IMAGE consortium cDNA clone isolated from the T-84 colonic epithelium cell line. AA100484 covers the bulk of Human STLK4 with its 3694 bp, which spans positions 146-3839 of SEQ ID NO:3. A second EST, AA297759, isolated from a Jurkat T cell cDNA library, spans positions 1-271 of the human STLK4 contig. The two ESTs overlap over a 126 bp stretch that has only one nucleotide discrepancy at position 149 (G in AA297759 and T in AA100484). A T at this position was chosen for the SEQ ID NO:3 based on sequence data generated from AA100484. The 5' 145 bp of human STLK4 contains three sequencing ambiguities (N's in SEQ ID NO:3) arising from sequence errors in the GenBank entry for AA297759. Three amino acid sequence ambiguities in the N-terminus of human STLK4 are present also in SEQ ID NO:7 as a consequence of the sequence inaccuracies from the EST entry.

The coding region of human STLK4 is 1242 bp long (2-1243), capable of encoding a 414 amino acid polypeptide, and is followed by a 2596 nucleotide 3' UTR (1244-3839). Human STLK4 ends in a polyadenylated stretch that has 18 adenines (3840-3857). A polyadenylation signal (AATAAA) is found between positions 3822-3827. Targeted-PCR cloning identified one rat orthologue of human STLK4, clone 135-31-19. In addition, one murine orthologue of human STLK4 was recognized in the EST database as AA117483. None of these orthologues add additional N-terminal sequence to the human STLK4.

The near full-length human STLK4 cDNA (SEQ ID NO.3) is 3857 bp long and consists of a 1242 bp ORF flanked by a 2596 bp 3' UTR (1244-3839) and an 18 nucleotide polyadenylated region. Polyadenylation signals (AATAAA) begin at positions

2181 and 3822. Since the coding region is open throughout the 5' extent of this sequence, this is apparently a partial cDNA clone lacking the N-terminal start methionine. A near full-length murine STLK4 cDNA is represented in the 1773 bp
5 EST AA117438. It extends an additional 21 nucleotides 5' of the human STLK4 consensus, but since its coding region is open throughout the 5' extent of the sequence, this is also probably a partial cDNA clone lacking the N-terminal start methionine.

10 Several EST fragments span the complete STLK3 sequence with AA297759 at the 5' end and AA100484 and others at the 3' end.

cDNA Cloning and Characterization of STLK5

15 The human STLK5 cDNA sequence is composed of four overlapping sequences, AI418298, 2R96-13-1, 3R25-45-3 and R46685. A human STLK5 clone, F07734, was first identified using a Smith-Waterman search of the EST database with SPS_sc (U33057) as a query.

20 AI418298 is an IMAGE consortium cDNA clone with an 895 bp insert.

PCR clone 2R96-13-1 was isolated from human brain sscDNA using primers 5'-CTCATCTGTACACACTTCATGG and 5'-GATTCCCACACTGTAGATGTC derived from F07734. 2R96-13-1
25 contains a 330 bp insert and an ORF of 330 bp (110 amino acids).

EST clone R46685 was identified using a Smith-Waterman search of the EST database with the C-terminus of SPS_sc (GB:U33057) as query. Sequence analysis of the 1047 bp
30 insert identified this EST to contain an ORF of 285 bp (95 amino acids) encoding the C-terminus of human STLK5.

PCR clone 3R25-45-3 was isolated from human fetal brain sscDNA using primers 5'- GGCCCTCGACTACATCCACCACAT and 5'-

CAACGAAACTAACACAGCATAAGG derived from 2R96-13-1 and R46685, respectively. 3R25-45-3 contains a 330 bp insert and an ORF of 750 bp (250 amino acids).

The full-length human STLK5 cDNA (SEQ ID NO:96) is 2110 bp long and consists of a 1119 bp ORF flanked by a 229 bp 5' UTR and a 762 bp 3' UTR. The sequence flanking the first ATG conforms to the Kozak consensus (*supra*) for an initiating methionine, and is believed to be the translational start site for STLK5.

Several EST fragments span the complete STLK5 sequence with AA297059 and F07734 at the 5' end and R46686 and F03423 and others at the 3' end.

STLK5 displays a 100% match over a 41 bp stretch (position 2-42, SEQ ID NO. 97) to a human CpG island repeat (Z61277).

cDNA Cloning and Characterization of STLK6

Human STLK6 was first identified in the translated EST database (AA219667) as a novel serine threonine kinase.

The partial human STLK6 cDNA (SEQ ID NO:98) is 2,001 bp long and consists of a 1,254 bp ORF flanked by a 75 bp 5' UTR and a 673 bp 3' UTR. The sequence flanking the first ATG conforms to the Kozak consensus (Kozak, M., Nucleic Acids Res. 15, 8125-8148 (1987)) for an initiating methionine, and is believed to be the translational start site for STLK6.

At the time of filing, inventors believe that STLK6 does not have any significant match in the nucleic acid database.

cDNA Cloning and Characterization of STLK7

Human STLK7 was first identified in the translated EST database (AA988954) as a novel serine threonine kinase. The

original clone was not available through public sources, so a PCR fragment amplified from the sequence of AA988954 yielded 5R54-21-2.

5 The partial human STLK7 cDNA (SEQ ID NO:100) is 311 bp long and consists of a 309 bp ORF. Since the coding region is open throughout the 5' and 3' extent of this sequence, this appears to be a partial cDNA clone lacking the N-terminal start methionine and C-terminal stop codon.

10 STLK7 shares 80% sequence identity to human SPAK (AF099989) over a 167 bp region and 50% nucleotide sequence identity to SLTK7 (SEQ ID NO. 101) over 391 nucleotides.

cDNA Cloning and Characterization of ZC1

15 The human ZC1 cDNA sequence is composed of two overlapping PCR clones, 3R25-24-2 and R65-12-2.

A human ZC1 clone, 125-33-5, was first identified from a PCR screen with degenerate oligos, TRK1 and TRK4, applied to sscDNA generated from human small airway epithelial cells (Clontech). Sequence analysis of the 503 bp insert 20 identified a 501 bp ORF (167 amino acids) with the potential to encode a novel human STK related to the *C. elegans* ZC504.4 gene product.

PCR clone 3R25-24-2 was isolated from human SNB19 glioblastoma sscDNA using primers 5'-ATGGCGAACGACTCTCCCGCAA 25 and 5'-ACACCAAAATCAACAAGTTCACCTC derived from the N-terminus of a murine orthologue of ZC1 (NIK, GB:U88984) and the original human ZC1 clone 125-33-5, respectively. 3R25-24-2 contains a 527 bp insert and an ORF of 519 bp (173 amino acids).

30 PCR clone R65-12-2 was isolated as follows: A Smith-Waterman search of the EST database with the *C. elegans* ZC504.4 gene (GB:Z50029) as a query identified a human EST (W81656) whose ORF is related to the *C. elegans* gene and

terminates in an identical residue (Trp). A primer was designed 3' to this stop codon (5'-AGTTACAAGGAATTCCAAGTTCT) and used in a PCR reaction with a primer derived from the original human ZC1 clone 125-33-5 (5'-

5 ATGAAGAGGAAGAAATCAAAC TG) using sscDNA from human SNB19 glioblastoma as a template. PCR clone R65-12-2 was identified and was found to contain a 3611 bp insert with a 3534 bp ORF encoding the C-terminal portion of human ZC1 (1178 amino acids).

10 The full-length human ZC1 cDNA (SEQ ID NO. 9) is 3798 bp long. Clone 3R25-24-2 spans positions 1-527, and clone R65-12-2 spans positions 188-3798. The overlap between these two clones exhibits 100% sequence identity. The human ZC1 contains a 3717 bp ORF (17-3723) flanked by a 6 bp 5' 15 UTR and a 75 bp (3724-3798) 3' UTR. No polyadenylation signal (AATAAA) or polyadenylated region are present in the 3'UTR. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for human ZC1.

20 Multiple EST fragments (W81656) match the 3' end of the human ZC1 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

25 cDNA Cloning and Characterization of ZC2

The human ZC2 cDNA sequence is composed of four overlapping PCR clones, G75-31-17, R65-24-6, 2R28-8-1, and R99-6-10.

30 A human ZC2 clone, G75-31-17, was first identified from a PCR screen with degenerate oligos, ROS1 (5'-GCNTTYGGNGARGTNTAYGARGG) and CCK4b (5'-GCTGGATCCYTCNGGNSWCATCCA), applied to sscDNA generated from the human HLT383 primary non-small cell lung cancer tissue.

Sequence analysis of the 492 bp insert identified a 492 ORF (164 amino acids) with the potential to encode a novel human STK related to the *C. elegans* ZC504.4 gene product.

PCR clone R99-6-10 was isolated as follows: A Smith-Waterman search of the EST database with *C. elegans* ZC504.4 gene (GB:Z50029) as a query identified two overlapping human EST fragments (AA115844 and R51245) whose ORFs were related to the *C. elegans* gene and terminate in an identical residue (Trp). A primer was designed 3' to the stop codon found in R51245 (5'-AGATGGACTGTACTGGGAGG) and used in a PCR reaction with a primer derived from AA115844 (5'-ACTTTGTGCAGCTCTGTGGG) using human fetal brain sscDNA as a template. PCR clone R99-6-10 was identified and was found to contain a 1095 bp insert with a 930 bp ORF encoding the C-terminal portion of human ZC2 (310 amino acids).

PCR clone R65-24-6 was isolated from human HT29 colon cancer cell line sscDNA using primers 5'-AAGGTTATGGATGTCACAGGG and 5'-AGATGGACTGTACTGGGAGG derived from G75-31-17 and R51245, respectively. The 3' primer used in this PCR reaction misprimed between positions 1634-1653 of this gene leading to the formation of a truncated product. R65-24-6 contains a 1593 bp insert and an ORF of 1593 bp (531 amino acids).

PCR clone 2R28-8-1 was isolated from human colon cancer cell line HT29 sscDNA using primers 5'-CTCACAAAGGTTGCCAACAGG and 5'-AGTCCCCACCAGAAAGGTTAC derived from R65-24-6 and R99-6-10, respectively. 2R28-8-1 contains a 1538 bp insert and an ORF of 1536 bp (512 amino acids).

The partial human ZC2 cDNA (SEQ ID NO. 10) is 4055 bp long. Clone G75-31-17 spans positions 1-492, clone R65-24-6 spans positions 58-1650, clone 2R28-8-1 spans positions 1466-3003 and clone R99-6-10 spans positions 2961-4055. The overlapping regions between these clones exhibit 100%

sequence identity except for a single guanine (G75-31-17) to adenosine (R65-24-6) mismatch at position 280 resulting in a Glu to Lys change. Based on the presence of an acidic residue in this position in human ZC1 and ZC3 and *C. elegans* ZC504.4, the sequence encoding the Glu is probably correct.

5 The human ZC2 gene contains a 3891 bp ORF (1-3891) flanked by 164 bp (3892-4055) 3' UTR. No polyadenylation signal (AATAAA) or polyadenylated region is present in the 3'UTR.

10 Multiple EST fragments (R51245) match the 3' end of the human ZC2 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

cDNA Cloning and Characterization of ZC3

15 The human ZC3 cDNA sequence is composed of four overlapping PCR clones, G75-30-30, 3R33-5-3, 3R19-17-6, and R99-43-11.

20 A human ZC3 clone, G75-30-30, was first identified from a PCR screen with degenerate oligos, ROS1 and CCK4b, applied to sscDNA generated from a human HLT370 primary non-small cell lung cancer tissue. Sequence analysis of the 492 bp insert identified a 492 ORF (164 amino acids) with the potential to encode a novel human STK related to the *C. elegans* ZC504.4 gene product.

25 PCR clone R99-43-11 was isolated as follows: A Smith-Waterman search of the EST database with the *C. elegans* ZC504.4 gene (GB:Z50029) as a query identified a human EST (R54563) whose ORF is related to the *C. elegans* gene and terminates in an identical residue (Trp). A primer was 30 designed 3' to the stop codon found in R54563 (5'-TCAGGGGTCAAGAGGTACG) and used in a PCR reaction with a primer derived from the 5' end of R54563 (5'-CCCAAACCCTACCACAAATTC) using sscDNA from human fetal brain

as a template. PCR clone R99-43-11 was identified and was found to contain a 719 bp insert with a 564 bp ORF encoding the C-terminal portion of human ZC3 (188 amino acids).

5 PCR clone 3R19-17-6 was isolated from human A549 lung cancer cell line sscDNA using primers 5'-
CCCCCGGGAAACGATGACCA and 5'-AGCCGCTGCCCTCCTACTGT derived from G75-30-30 and R99-43-11, respectively. The 3' primer used in this PCR reaction misprimed leading to the formation of a truncated product. 3R19-17-6 contains a 1172 bp insert
10 and an ORF of 1170 bp (390 amino acids).

PCR clone 3R33-5-3 was isolated from human A549 lung cancer cell line sscDNA using primers 5'-
ACCGAACATGCCACCTACTAC and 5'-CTCGACGTCGTGGACCACC derived from G75-30-30 and 3R19-17-6, respectively. 3R33-5-3
15 contains a 2465 bp insert and an ORF of 2463 bp (821 amino acids).

20 The full-length human ZC3 cDNA (SEQ ID NO. 11) is 4133 bp long. Clone G75-30-30 spans positions 1-483, clone 3R33-5-3 spans positions 134-2598, clone 3R19-17-6 spans positions 2356-3512 and clone R99-43-11 spans positions 3415-4133. The overlaps between these clones exhibit 100% sequence identity. The human ZC3 gene contains a 3978 bp ORF (1-3978) flanked by a 152 bp 3' UTR (3979-4133). No polyadenylation signal (AATAAA) or polyadenylated region is
25 present in the 3'UTR.

Multiple EST fragments (R54563) match the 3'end of the human ZC3 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

cDNA Cloning and Characterization of ZC4

5

The human ZC4 cDNA sequence, represented by PCR fragment 3R25-27-1, was first identified in the human genomic cosmid 82J11 (GB:Z833850) containing exon sequences that displayed strong homology to the ZC504.4 *C. elegans* gene.

10

PCR clone 3R25-27-1 was isolated from human fetal liver ssDNA and primers 5'-CAATGTTAACCCACTCTATGTCTC and 5'-AGTTTGCGATGTTTCTTTTC derived from a potential ORF (positions 25729-25852) from the 82J11 cosmid and from an EST (R98571) encoding the C-terminus of the human ZC4 gene, respectively.

15

The partial human ZC4 cDNA (SEQ ID NO.12) is 1459 bp long and consists of a 1047 bp ORF (2-1048) flanked by a 411 bp (1049-1459) 3'UTR region. No polyadenylation signal (AATAAA) or polyadenylated region is present in the 3'UTR.

20

The N-terminal coding sequence for ZC4_h was extended by building a contiguous DNA sequence of 233,137 bp containing Z83850 and four other sequences: CU84B10 and CU230B10 (from the Sanger Human Genome Sequencing Project, <http://www.sanger.ac.uk/HGP/>) and Z97356 and Z69734 (available from the National Institute for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/Entrez/nucleotide.html>). The position of each sequence in the contig is represented in the table below.

25

Accession	Length	Start	End
CU84B10	43273		43273
Z97356	21848	43170	65018
Z69734	37077	63077	100149
CU230B10	11841	8841	100256
Z83850	132981	100155	233137

30

Sequences in ZC4 genomic contig.

The 233,137 bp contig was analyzed for exons using the programs FGENES 1.5 and FGENESH, human gene structure prediction software available from The Sanger Centre (<http://genomic.sanger.ac.uk/qf/qf.shtml>).

5 The resulting human ZC4 coding sequence (SEQ ID NO:104) is 3,681 bp long (excluding the stop codon) and encodes for a STE20 kinase of 1227 amino acids.

cDNA Cloning and Characterization of KHS2

10 The human KHS2 cDNA sequence is composed of four overlapping clones, 3R25-51-2, 3R16-34-2, 3R16-31-2, and T79916.

15 A human KHS2 clone, AA250855, was first identified using a Smith-Waterman search of the EST database with KHS1 (GB:U77129) as a query. Sequence analysis of the 1112 bp insert identified a 618 bp ORF (206 amino acids) with the potential to encode a novel STK related to the human KHS1 gene product. Using AA250855 as a query, a second EST (AA446022) was found whose sequence was shown to contain the 20 initiator methionine for human KHS2 based on a comparison with KHS1.

25 PCR clone 3R25-51-2 was isolated from human testicular cancer sscDNA using primers 5'-CCGCCATGAACCCGGCTT and 5'-CGATTGCCAAAGACCGTGTCA derived from AA446022 and AA250855, respectively. 3R25-51-2 contains an 850 bp insert and an ORF of 849 bp (283 amino acids).

EST clone, T79916, was identified using a Smith-Waterman search of the EST database with the C-terminus of KHS1 (GB:U77129) as a query. Sequence analysis of the 2107 bp insert identified this EST to contain an ORF of 345 bp (115 amino acids disrupted by a single stop codon) encoding the C-terminus of human KHS2, followed by 1762 bp 3'UTR.

PCR clone 3R16-34-2 was isolated from human testis sscDNA using primers 5'-AGAAGTTGCAGCTGTTGAGAGGA and 5'-TATGGCCCGTGTAAAGGATTTC derived from AA250885 and T79916, respectively. 3R16-34-2 contains an 1516 bp insert and an ORF of 1128 bp (376 amino acids).

5 PCR clone 3R16-31-2 was isolated from normal human colon sscDNA using primers 5'- GTGCCAGAAGTGTGTTGTAA and 5'-TATGGCCCGTGTAAAGGATTTC derived from EST T79916. 3R16-31-2 contains a 728 bp insert and an ORF of 669 bp (223 amino acids). This clone lacked the stop codon present within EST 10 T79916 (position 2662 in the KHS2 sequence).

10 The full-length human KHS2 cDNA (SEQ ID NO.17) is 4023 bp long. Clone 3R25-51-2 spans positions 1-855, clone AA250885 spans positions 336-923, clone 3R16-34-2 spans 15 positions 545-2061, and clone T79916 spans positions 1917-4023. The overlapping regions between these clones exhibit 20 100% sequence identity, except for 4 nucleotide differences, two of which are silent, a third corrects the internal stop codon at position 2662, and the fourth at position 247 (T to C change) results in a Pro to Leu change. The human KHS2 cDNA contains a 2682 bp ORF (6-2687) flanked by a 5 bp (1-5) 25 5'UTR and a 1336 bp (2688-4023) 3' UTR. A potential polyadenylation signal (AATAAA) is found at positions 4008-4013. No polyadenylated region is present in the 3'UTR. The sequence flanking the first ATG is in a poor context for translational initiation, however, a 134 bp 5'UTR sequence from EST AA446022 did not reveal any additional ATG's and displayed two in-frame stop codons 5' to the putative start ATG for human KHS2.

30 Multiple EST fragments match the 5'end (AA446022) as well as the 3' end (R37625) of the human KHS2 gene.

161

cDNA Cloning and Characterization of SULU1

The human SULU1 cDNA sequence is composed of three overlapping clones, N40091, 2R90-1-1 and R90907.

A human SULU1 clone, N40091, was first identified using a Smith-Waterman search of the EST database with the *C. elegans* SULU gene (GB: U32275) as a query. Sequence analysis of the 1321 bp insert identified a 906 bp ORF (302 amino acids) with the potential to encode a novel human STK related to the *C. elegans* SULU gene product.

EST clone R90907 was first identified using a Smith-Waterman search of the EST database with the 3' end of the *C. elegans* SULU gene (GB: U32275) as a query. Sequence analysis of the 1647 bp insert identified a 578 bp ORF (192 amino acids) with the potential to encode the C-terminus of the human SULU1 gene product.

PCR clone 2R90-1-1 was isolated from human HT29 colon cancer cell sscDNA using primers 5'- TATTGAATTGGCGGAACCGGAAG and 5'- TTGTTTGTGCTCATTCTTGGAG derived from N40091 and R90907, respectively. 2R90-1-1 contains a 1625 bp insert and an ORF of 1623 bp (541 amino acids).

The full-length human SULU1 cDNA (SEQ ID NO.19) is 4177 bp long. Clone N40091 spans positions 1-1321, clone 2R90-1-1 spans positions 1048-2671, and clone R90907 spans positions 2531-4177. The overlapping regions between these clones exhibit 100% sequence identity. The human SULU1 cDNA contains a 2694 bp ORF (415-3108) flanked by a 414 bp (1-414) 5'UTR and a 1069 bp (3109-4177) 3' UTR followed by a 19 nucleotide polyadenylated region. A potential polyadenylation signal (AATAAA) is found at positions 4164-4169. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for human SULU1.

Multiple EST fragments match the 5' end (N27153) as well as the 3' end (R90908) of the human SULU1 gene.

cDNA Cloning and Characterization of Murine SULU3

5 The murine SULU3 cDNA sequence is represented by PCR fragment 2R92-1-6.

A murine SULU3 clone, G83-4-5, was first identified from a PCR screen with degenerate oligos, CCK4c and CCK4b, applied to sscDNA generated from murine day-12 embryos.

10 Sequence analysis of the 473 bp insert identified a 471 ORF (157 amino acids) with the potential to encode a novel human STK related to the *C. elegans* SULU gene (GB: U32275) product. The antisense strand of G83-4-5 is identical at the nucleic acid level to the 5'UTR of the murine ets1 protooncogenic transcription factor (GB:X53953). This homology is likely the result of a cloning artifact attached to the 5'-end of the database entry for murine ets1.

15 PCR clone 3R19-17-6 was isolated from human A549 cell sscDNA using primers 5'-CCCCCGGGAAACGATGACCA and 5'-AGCCGCTGCCCTCCTACTGT derived from G75-30-30 and R99-43-11, respectively. The 3' primer used in this PCR reaction misprimed leading to the formation of a truncated product. 3R19-17-6 contains a 1172 bp insert and an ORF of 1170 bp (390 amino acids).

20 PCR clone 2R92-1-6 was isolated from murine d8 embryo sscDNA using primers 5'-ACCGAACATGCCACCTACTAC and 5'-GATTGCTTGTGCTATTCTTG derived from the 5' UTR of the ets1 gene and the human EST AA234623, respectively. The latter (shown herein) encodes the C-terminus of human SULU3. 2R92-30 1-6 contains a 2249 bp insert and an ORF of 2244 bp (748 amino acids).

The partial murine SULU3 cDNA (SEQ ID NO.21) is 2249 bp long and consists of a 2244 bp ORF (6-2249) flanked by a 5

bp (1-5) 5'UTR. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start site for murine SULU3.

5 One EST fragment (AA446022) matches the 3' end of the partial murine SULU3 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

10 cDNA Cloning and Characterization of Human SULU3

The human SULU3 cDNA sequence is composed of two overlapping clones, 2R90-22-1 and AA234623.

15 A human SULU3 clone, AA234623, was first identified using a Smith-Waterman search of the EST database with the *C. elegans* SULU gene (GB: U32275) as a query. Sequence analysis of the 2652 bp insert identified a 1185 bp ORF (395 amino acids) with the potential to encode the C-terminus of a novel human STK related to the *C. elegans* SULU gene product.

20 PCR clone 2R90-22-1 was isolated from human SKMel128 melanoma cell line sscDNA using primers 5'-TATTGAATTGGCGGAACCGGAAG and 5'- TTGTTCTAACAGAGTGCCCTCCG derived from the murine SULU3 2R92-1-6 clone and from AA234623, respectively. 2R92-1-6 contains a 1897 bp insert and an ORF of 1896 bp (632 amino acids).

25 The partial human SULU3 cDNA (SEQ ID NO.20) is 3824 bp long. Clone 2R90-22-1 spans positions 1-1897 and clone AA234623 spans positions 1173. The overlapping region between these clones exhibits 100% sequence identity. The human SULU3 cDNA contains a 2358 bp ORF (2-2359) flanked by a 1465 bp (2360-3824) 3'UTR followed by a 19 nucleotide polyadenylated region. A potential polyadenylation signal (AATAAA) is found at positions 2602-2607. Since the coding

region is open throughout the 5' extent of this sequence, this is apparently a partial cDNA clone lacking the N-terminal start methionine.

5 Multiple EST fragments (R02283) match the 3' end of the human SULU3 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

cDNA Cloning and Characterization of GEK2

10 The human GEK2 cDNA sequence is composed of three overlapping clones, AA459448, 3R25-48-1 and GEK2_h#3.

15 A human GEK2 clone, AA459448, was first identified using a Smith-Waterman search of the EST database with the human SLK gene (GB: AB002804) as a query. Sequence analysis of the 1286 bp insert identified a 1227 bp ORF (409 amino acids) with the potential to encode the N-terminus of a novel human STK related to the human SLK gene product. An additional Smith-Waterman search using the C-terminus of the SLK gene as a query yielded three additional EST's, 20 AA323687, AA380492 and AA168869, that encode the C-terminal region of human GEK2.

25 PCR clone 2R98-41-17 was isolated from human testis sscDNA using primers 5'- AAGACCATGCCGTGCGCCG and 5'- ATTCCCTTCAGGTTCTGGTTATGG derived from AA323687 and from AA380492, respectively. 2R98-41-17 contains a 851 bp insert and an ORF of 849 bp (283 amino acids).

30 PCR clone GEK2_h#3 was isolated from human sscDNA made from the H23 tumor cell line using primers 5'- GCAGCAAAGTGGAGAAGATGG and 5'- GGAAGCATCCCCAGAGCTGTAG derived from the sequence of clone 3R25-48-1 and from the 3' end of murine LOK (GB:D89728), respectively. GEK2_h#3 contains a 1042bp insert and an ORF of 1041 bp (347 amino acids).

The full-length human GEK2 cDNA (SEQ ID NO:106) is 2962 bp long. Clone AA459448 spans positions 1-1286, clone 3R25-48-1 spans positions 1100-2449 and clone GEK2_h#3 spans positions 1920-2962. The overlapping regions between these 5 clones exhibit 100% sequence identity.

The human GEK2 cDNA contains a 2904 bp ORF (59-2962) flanked by a 58 bp (1-58) 5'UTR. The sequence flanking the first ATG conforms to the Kozak consensus for an initiating methionine, and is believed to be the translational start 10 site for human GEK2.

Multiple EST fragments (AA465671) match the 5'end of the sequence, but only one (AA380492) matches the 3' end of the human GEK2 gene.

15 cDNA Cloning and Characterization of PAK4

The human PAK4 cDNA sequence is represented by clone SNB2#1.

A human PAK4 clone, R88460, was first identified using a Smith-Waterman search of the EST database with the human 20 PAK gene (GB: U24152) as a query. Sequence analysis of the 2332 bp insert identified a 930 bp ORF (310 amino acids) with the potential to encode the C-terminus of a novel human STK related to the human PAK gene product.

25 cDNA clone SNB2#1 was isolated from human glioblastoma cell line SNB75 cDNA library using a probe derived from R88460. SNB2#1 contains a 3604 bp insert and an ORF of 2043 bp (681 amino acids).

The full-length human PAK4 cDNA (SEQ ID NO.27) is 3604 bp long and consists of a 2043 bp ORF (143-2185) flanked by 30 a 142 bp (1-142) 5'UTR and a 1419 3' UTR followed by a 22 nucleotide polyadenylated region. A potential polyadenylation signal (AATTAAA) is found at positions 3582-3588. The sequence flanking the first ATG conforms to the

166

Kozak consensus for an initiating methionine, and is believed to be the translational start site for human PAK4. The 3' UTR of the PAK4 gene contains a GT dinucleotide repeat prone to undergo expansion based on the number of repeats found in clones SNB#1 and R88460, 32 and 23, respectively. Several neurologic disorders have been correlated with the expansion of di- or tri-nucleotide repeats similar to those found in the PAK4 sequence, suggesting PAK 4 may also be a disease target and that this repeat in its 3'UTR may serve as a diagnostic marker.

Multiple EST fragments (AA535791) match the 3'end of the human PAK4 gene, but at the time of filing, the inventors believe that none exist in GenBank or the EST database that match its 5' end.

15

cDNA Cloning and Characterization of PAK5

The full-length human PAK5 cDNA sequence is composed of two overlapping clones, H450#1-1 and SNB8#5.

A human PAK5 clone, R18825, was first identified using a Smith-Waterman search of the EST database with the human PAK4 gene as a query. Sequence analysis of the 1248 bp insert identified a 420 bp ORF (140 amino acids) with the potential to encode the C-terminus of a novel human STK related to the human PAK4 gene product.

25

cDNA clone SNB8#5 was isolated from human SNB75 cDNA library using a probe derived from R18825. SNB2#1 contains a 2028 bp insert and an ORF of 1194 bp (398 amino acids).

30

The partial human PAK5 cDNA (SEQ ID NO.28) is 2028 bp long and consists of a 1194 bp ORF (2-1195) flanked by an 833 bp (1196-2028) 3'UTR followed by a 22 nucleotide polyadenylated region. A potential polyadenylation signal (AATTAAA) is found at positions 2004-2010. Since the coding region is open throughout the 5' extent of this sequence,

this is apparently a partial cDNA clone lacking the N-terminal start methionine.

Clone H460#1-1 was isolated from a human lung H460 cDNA library using a probe derived from the partial SNB2#1 cDNA clone described above. Sequence analysis of the 2526 bp insert identified a 1773 bp ORF (592 amino acids) with the potential to encode a full-length PAK5.

The human PAK5 cDNA (SEQ ID NO:102) is 2,806 bp long and consists of a 1,773 bp ORF flanked by a 201 bp 5' UTR and a 833 bp 3' UTR. The sequence flanking the first ATG conforms to the Kozak consensus (Kozak, M., Nucleic Acids Res. 15, 8125-8148 (1987)) for an initiating methionine, and is believed to be the translational start site for PAK5.

PAK5 shares 99% sequence identity over 2795 bp to a recent database entry, AF005046. These sequences are presumed to be from the same gene, with minor polymorphic variations.

EXAMPLE 2: Expression Analysis of Mammalian STE20-related Protein Kinases

Materials and Methods

Northern blot analysis

Northern blots were prepared by running 10 µg total RNA isolated from 60 human tumor cell lines (HOP-92, EKVX, NCI-H23, NCI-H226, NCI-H322M, NCI-H460, NCI-H522, A549, HOP-62, OVCAR-3, OVCAR-4, OVCAR-5, OVCAR-8, IGROV1, SK-OV-3, SNB-19, SNB-75, U251, SF-268, SF-295, SF-539, CCRF-CEM, K-562, MOLT-4, HL-60, RPMI 8226, SR, DU-145, PC-3, HT-29, HCC-2998, HCT-116, SW620, Colo 205, HTC15, KM-12, UO-31, SN12C, A498, CaK1, RXF-393, ACHN, 786-0, TK-10, LOX IMVI, Malme-3M, SK-MEL-2, SK-MEL-5, SK-MEL-28, UACC-62, UACC-257, M14, MCF-7, MCF-7/ADR RES, Hs578T, MDA-MB-231, MDA-MB-435, MDA-N, BT-549, T47D), from 22 human adult tissues (thymus, lung,

duodenum, colon, testis, brain, cerebellum, cortex, salivary gland, liver, pancreas, kidney, spleen, stomach, uterus, prostate, skeletal muscle, placenta, mammary gland, bladder, lymph node, adipose tissue), and 2 human fetal normal tissues (fetal liver, fetal brain), on a denaturing formaldehyde 1.2% agarose gel and transferring to nylon membranes.

Filters were hybridized with random primed [$\alpha^{32}\text{P}$]dCTP-labeled probes synthesized from the inserts of several of the STE20-related kinase genes. Hybridization was performed at 42 °C overnight in 6X SSC, 0.1% SDS, 1X Denhardt's solution, 100 µg/mL denatured herring sperm DNA with 1-2 x 10⁶ cpm/mL of ³²P-labeled DNA probes. The filters were washed in 0.1X SSC/0.1% SDS, 65 °C, and exposed on a Molecular Dynamics phosphorimager.

Quantitative PCR analysis

RNA was isolated from a variety of normal human tissues and cell lines. Single stranded cDNA was synthesized from 10 µg of each RNA as described above using the Superscript Preamplification System (GibcoBRL). These single strand templates were then used in a 25 cycle PCR reaction with primers specific to each clone. Reaction products were electrophoresed on 2% agarose gels, stained with ethidium bromide and photographed on a UV light box. The relative intensity of the STK-specific bands were estimated for each sample.

DNA Array Based Expression Analysis

Plasmid DNA array blots were prepared by loading 0.5 µg denatured plasmid for each STE20-related kinase on a nylon membrane. The [$\alpha^{32}\text{P}$]dCTP labeled single stranded DNA probes were synthesized from the total RNA isolated from several

human immune tissue sources or tumor cells (thymus, dendrocytes, mast cells, monocytes, B cells (primary, Jurkat, RPMI8226, SR), T cells (CD8/CD4+, TH1, TH2, CEM, MOLT4), K562 (megakaryocytes). Hybridization was performed
5 at 42 °C for 16 hours in 6X SSC, 0.1% SDS, 1X Denhardt's solution, 100 µg/mL denatured herring sperm DNA with 10⁶ cpm/mL of [$\alpha^{32}\text{P}$]dCTP labeled single stranded probe. The filters were washed in 0.1X SSC/0.1% SDS, 65 °C, and exposed
10 for quantitative analysis on a Molecular Dynamics phosphorimager.

RESULTS

Distribution of STE20-Related Gene Transcripts in Normal Tissues and Tumor Cell Lines

15 ZC1, ZC2, and ZC3 RNA expression was analyzed by quantitative PCR from multiple human normal tissues, cultured primary epithelial and endothelial cells, and tumor cell lines. The results are summarized in Tables 1 and 2, with relative expression values ranging from 0
20 (undetectable) to 23 (very strong). An "x" refers to sample not tested. ZC1, ZC2, and ZC3 were all expressed at very low levels in most normal human tissues, however ZC1 and ZC2 were more abundant in cultured epithelial cells and ZC3 in normal kidney and breast tissue.

25 Expression of these 3 genes was also examined in a panel of human tumor cell lines representing a diverse sampling of tumor types (Table 2). ZC1 and ZC2 showed strong expression in cell lines from most melanomas and renal tumors and from some non-small cell lung cancers and colon tumors. ZC3 expression was consistently lower in the tumor cell lines except for high expression in most breast cancers and leukemias. The robust overexpression ZC1, ZC2,
30

and ZC3 in tumor cells versus normal tissues may provide an attractive target for oncology drug development.

Expression of all the novel STE20-related kinases was examined in a panel of human immune tissues/cells by hybridization to a DNA array blot containing plasmids encoding each of these genes. STLK2 was broadly expressed in all 14 immune samples, whereas STLK4 and PAK4 were highly expressed in a subset of 6-7 of the samples (Table 3).

Several other kinases (SULU3, ZC4, KHS2) had more restricted expression, while others were expressed in only a single immune source (STLK3, thymus; ZC1, dendrocytes; ZC3, monocytes; PAK5, mast cells and MOLT4), and several more were absent from all the immune sources assayed (GEK2, SULU1, ZC2, STLK5). These expression patterns were quite distinct among members of the same subfamily (*i.e.*, ZC1, ZC2, ZC3 and ZC4, or PAK1, PAK2, PAK3, PAK4, PAK5). This analysis suggests that some of these kinases may be candidate targets for various immune disorders, and that some, which are more broadly expressed, may mediate functions vital to the basic biology of most proliferating cells.

TABLE 1

ZC1, ZC2 and ZC3 Expression
in Normal Human Tissues and Cells

Sample		ZC1	ZC2	ZC3
NORMAL				
Brain	Tiss	2.8	0.6	0.9
Duod	Tiss	3.8	1.5	0.3
Heart	Tiss	1.2	0.3	0.0
Kidney	Tiss	0.7	0.0	7.0
Lung	Tiss	1.6	0.2	0.0
Pancreas	Tiss	2.0	0.4	2.5

Placenta	Tiss	1.4	0.0	0.0
Sal gl.	Tiss	3.0	0.3	3.2
Sk mus.	Tiss	2.3	0.1	0.1
Spleen	Tiss	0.4	0.0	x
Stomach	Tiss	0.8	0.0	0.0
Thymus	Tiss	3.5	0.4	1.5
Cereb	Tiss	2.8	1.1	4.4
Liver	Tiss	1.8	0.0	0.4
Uterus	Tiss	1.6	0.0	1.4
Prostate	Tiss	1.4	0.0	1.6
Testis	Tiss	x	x	5.8
f Brain	Tiss	x	x	3.1
Mam gl	Tiss	x	x	7.2
HCAEC	ENDO	1.0	0.0	0.0
HMVEC-d	ENDO	0.7	0.0	0.4
HMVEC-L	ENDO	2.2	1.6	1.8
HPAEC	ENDO	9.3	5.3	6.4
HMEC	EPI	4.1	2.3	1.9
RPTEC	EPI	3.6	2.2	0.2
HRCE	EPI	5.3	3.5	1.3
HSAE	EPI	0.9	3.3	4.8

TABLE 2

ZC1, ZC2 and ZC3 Expression in Tumor Cell lLnes

Sample	Origin	ZC1	ZC2	ZC3
HOP-92	Lung	9.3	7.2	3.3
EKVK	Lung	10.7	3.7	3.5
NCI-H23	Lung	5.8	6.3	4.1
NCI-H226	Lung	6.5	6.8	3.3

Sample	Origin	ZC1	ZC2	ZC3
HCC-2998	Colon	2.4	3.8	3.0
HCT 116	Colon	2.2	2.1	5.4
SW-620	Colon	7.8	12.1	3.1
COLO 205	Colon	9.1	16.2	3.0

NCI-H322M	Lung	3.5	5.8	4.9
NCI-H460	Lung	4.5	3.7	2.9
NCI-H522	Lung	4.7	3.3	4.6
A549/ATCC	Lung	3.8	3.6	4.1
HOP-62	Lung	4.3	3.8	4.2
OVCAR-3	Ovary	2.9	3.1	1.5
OVCAR-4	Ovary	3.3	1.0	3.8
OVCAR-5	Ovary	2.6	3.6	2.2
OVCAR-8	Ovary	3.6	2.0	4.7
IGROV1	Ovary	3.8	1.7	3.2
SK-OV-3	Ovary	4.9	0.0	3.5
SNB-19	CNS	5.1	5.4	4.2
SNB-75	CNS	2.5	0.9	0.7
U251	CNS	1.5	1.2	0.6
SF-268	CNS	5.8	2.7	3.0
SF-295	CNS	6.4	1.1	3.2
SF-539	CNS	5.1	2.9	4.3
CCRF-CEM	Leuk	3.4	2.7	3.1
K-562	Leuk	4.1	6.3	4.3
MOLT-4	Leuk	7.1	3.4	4.2
HL-60	Leuk	x	x	0.4
RPMI 8226	Leuk	0.5	0.2	1.4
SR	Leuk	3.5	7.2	5.4
DU-145	Pro	x	x	3.4
PC-3	Pro	x	x	3.4
HT-29	Colon	2.4	5.9	6.6

HCT-15	Colon	13.8	4.9	2.5
KM-12	Colon	7.0	13.2	3.1
UO-31	Colon	10.4	10.6	0.9
SN12C	Renal	8.1	3.4	2.8
A498	Renal	6.2	3.1	2.9
Caki-1	Renal	9.2	14.4	2.3
RXF 393	Renal	10.6	4.8	2.8
ACHN	Renal	9.3	6.0	3.9
786-0	Renal	8.8	15.6	5.6
TK-10	Renal	20.9	21.2	5.0
LOX IMVI	Mel	2.3	2.4	3.3
Malme-3M	Mel	x	x	2.2
SK-MEL-2	Mel	15.7	14.1	2.9
SK-MEL-5	Mel	7.9	7.0	0.0
SK-MEL-28	Mel	16.5	23.1	0.0
UACC-62	Mel	12.1	18.3	5.3
UACC-257	Mel	10.8	9.4	6.2
M14	Mel	4.4	0.9	7.9
MCF7	Breast	4.8	1.3	7.7
MCF-7/ADR	Breast	8.8	3.4	7.7
Hs 578T	Breast	6.9	2.6	5.7
MDA-MB-231	Breast	5.7	1.9	6.4
MDA-MB-435	Breast	4.8	6.7	9.1
MDA-N	Breast	7.3	6.3	9.1
BT-549	Breast	3.6	1.9	8.0
T-47D	Breast	0.4	12.3	9.3

Table 3: STE20-related kinase expression in a human immune panel

KINASE	thymus	Dendro- cytes	Mast cells	Mono- cytes	B cells	CD8+ CD4+	TH1	TH2
GEK2	350	350	350	350	350	350	350	350
SULU1	350	350	350	350	350	350	350	350
SULU3	350	350	350	350	12149	350	5115	350
STLK2	117770	13771	27620	92036	18305	39109	5408	3564
STLK3	8624	350	350	350	350	350	350	350
STLK4	8524	350	350	350	350	8685	5642	350
STLK5	xxx	xxx	xxx	xxx	350	350	350	xxx
ZC1	350	3377	350	350	350	350	350	350
ZC2	350	350	350	350	350	350	350	350
ZC3	350	350	350	20156	350	350	350	350
ZC4	xxx	xxx	xxx	xxx	350	350	350	xxx
KHS2	8766	2508	350	56575	350	350	350	350
PAK4	32658	7684	3729	100948	350	350	350	1604
PAK5	350	350	4905	350	350	350	350	350

KINASE	CEM (T cell)	MOLT4 (T cell)	JURKAT (B cell)	RPMI8226 (B cell)	SR (B cell)	K562 (MO)
GEK2	350	350	350	350	350	350
SULU1	350	350	350	350	350	350
SULU3	350	350	350	350	350	350
STLK2	47236	53262	47605	22560	65936	30390
STLK3	350	350	350	350	350	350
STLK4	3648	350	26772	1570	350	350
STLK5	350	350	350	xxx	350	350
ZC1	350	350	350	350	350	350

174

ZC2	350	350	350	350	350	350
ZC3	350	350	350	350	350	350
ZC4	1094	7813	14945	xxx	350	6385
KHS2	350	350	350	350	350	350
PAK4	350	10246	350	3229	350	350
PAK5	350	12672	350	350	350	350

Transcript size from Northern data

Kinase	(kb)
STLK2	3.8
STLK4	5.0
ZC1	6.9/4.7
ZC2	6.0/8.0
ZC4	5
KHS2	4.4
SULU1	4.5
SULU3	10.0
GEK2	5.5
PAK4	4.8
PAK5	3.5

STLK2 is widely expressed; the highest expression levels were found in placenta, spleen and PBL.

STLK4 is also widely expressed in normal tissues including heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon, and peripheral blood lymphocytes. STLK4 was also detected in Jurkat T cells.

ZC1 is highly overexpressed in the following human cancer cell lines: HOP-92, EKVX, NCI-H23, NCI-H226, NCI-H322M, NCI-H522, A549, HOP-62 (lung); OVCAR-3, OVCAR-4, OVCAR-5 (ovary); SNB-19, U251, SF-268, SF-295, SF-539 (CNS); K-562, RPMI-8226 (leukemia); DU-145, PC-3 (prostate); HT-29, HCC-2998, HCT-116, SW620, COLO-205, HCT-15, KM-12 (colon); UO-31, CAKi-1, RXF-393, 786-0, TK-10 (renal); LOXIMVI, Malme-3M, SK-MEL-2, SK-MEL-28, UACC-62, UACC-257, M14

(melanoma); and MCF-7, MCF-7/ADR, HIS 578T, MDA-MB-231, MDA-MB-431, MDA-N, BT-549, T-47D (breast).

ZC2 is expressed in brain and testis. It is highly overexpressed in the following human cancer cell lines: TK-5 10 (renal); SK-MEL-28, UACC-62 (melanoma); T47D (breast).

Moderate expression in HOP92 (lung); OVCAR4, IGROV1 (ovary); DNB75, U251 (brain); K-562 (leukemia); and COLO205 (colon).

SULU1 is overexpressed in the following human cancer 10 cell lines: HOP-92, EKX, NCI-H23, NCI-H226, NCI-H322M, NCI-H522, A549, HOP-62 (lung); OVCAR-3, OVCAR-4, OVCAR-5, SK-OV-3 (ovary); SNB-19, U251, SF-268, SF-295, SF-539 (CNS); K-562, RPMI-8226 (leukemia); DU-145, PC-3 (prostate); HT-29, HCC-2998, HCT-116, SW620, COLO-205, HCT-15, KM-12 (colon); 15 UO-31, CAKi-1, RXF-393, 786-0, TK-10 (renal); LOX, IMVI, Malme-3M, SK-MEL-2, SK-MEL-28, UACC-62, UACC-257, M14 (melanoma); MCF-7, MCF-7/ADR, HIS 578T, MDA-MB-231, MDA-MB-431, MDA-N, BT-549, T-47D (breast)

SULU3 showed a broad pattern of expression in the 20 normal tissue panel of RNAs.

GEK2 was expressed in spleen, thymus and testis. Expression was high in the cell lines RBL-2H3 and H441.

PAK4 was expressed in the normal tissues: brain, testis and prostate, and in the human cancer cell lines: HNCI-H23 25 (lung); OVCAR-3 (ovary); SNB-19, U251 (CNS); RPMI-8226 (leukemia); DU-145 (prostate); COLO-205, HCT-15 (colon).

PAK5 showed weak expression levels in the normal 30 tissues: brain, testes, bladder, colon, adrenal medulla, spleen, fetal liver, breast, cerebral cortex, cerebellum, thymus, salivary gland, lung, stomach, duodenum, uterus, prostate, skeletal muscle and placenta. PAK5 was overexpressed in the human cancer cell lines: HOP-92, EKX, NCI-H23, NCI-H226, NCI-H322M, NCI-H522, A549, HOP-62 (lung);

OVCAR-3, OVCAR-4, OVCAR-5, SK-OV-3 (ovary); SNB-19, U251,
SF-268, SF-295, SF-539 (CNS); K-562, RPMI-8226 (leukemia);
DU-145, PC-3 (prostate); HT-29, HCC-2998, HCT-116, SW620,.
5 COLO-205, HCT-15, KM-12 (colon); UO-31, CAKi-1, RXF-393,
786-0, TK-10 (renal); LOXIMVI, Malme-3M, SK-MEL-2, SK-MEL-
28, UACC-62, UACC-257, M14 (melanoma); MCF-7, MCF-7/ADR, HIS
578T, MDA-MB-231, MDA-MB-431, MDA-N, BT-549, T-47D (breast).

10 EXAMPLE 3: STE20-related Protein Kinase Gene Expression
Vector Construction

Materials and Methods

Expression Vector Construction

Several expression constructs were generated for some
of the human STE20-related cDNAs including: a) full-length
15 clones in a pCDNA expression vector; b) a GST-fusion
construct containing the catalytic domain of the novel
STE20-related kinase fused to the C-terminal end of a GST
expression cassette; and c) a full-length clone containing a
Lys to Ala (K to A) mutation at the predicted ATP binding
20 site within the kinase domain, inserted in the pCDNA vector.

The "K to A" mutants of the STE20-related kinase might
function as dominant negative constructs, and will be used to
elucidate the function of these novel STKs.

25

RESULTS

Constructs for ZC1, ZC2, ZC3, SULU1, SULU3, PAK4 and PAK5
have been generated.

Numerous additional constructs have been generated for
the various STE20-subfamily kinases, including full length,
30 kinase inactive and tagged versions. In addition, the
following three constructs were designed for specific
applications based on their unique domain structure:

Construct 1: SULU1-coiled-coil2

Vector: pGEX-4T

Insert: Coiled-coil2

Sequence: Amino acids 752-898

5 Purpose: phage display

Result: Interacts with GEK2 CC1

Construct 2: SULU3-coiled-coil2

Vector: pGEX4T

10 Insert: coiled-coil 2 domain fused to GST

Sequence range of insert: amino acids 802-898 of SEQ

Purpose: phage display

Result: Interacts with coiled-coiled region of human SLK

15 Construct 3: PAK5 Dominant Negative

Vector: pCAN5

Insert: Full length coding sequence of human PAK5 containing the following mutation: K350,351A (Lys at aa positions 350 and 351 changed to Ala).

20 Purpose: to determine role of human PAK5 kinase activity in cell growth and transformation.

Result: Interferes with Ras transformation.

25 EXAMPLE 4: Generation of Specific Immunoreagents to STE20-Related Protein Kinases

Materials and Methods

Specific immunoreagents were raised in rabbits against KLH- or MAP-conjugated synthetic peptides corresponding to the human STE20-related kinases. C-terminal peptides were conjugated to KLH with glutaraldehyde, leaving a free C-terminus. Internal peptides were MAP-conjugated with a blocked N-terminus. Additional immunoreagents can also be generated by immunizing rabbits with the bacterially expressed

GST-fusion proteins containing the cytoplasmic domains of each novel STK.

The various immune sera are first tested for reactivity and selectivity to recombinant protein, prior to testing for endogenous sources.

Western blots

Proteins in SDS PAGE are transferred to immobilon membrane. The washing buffer is PBST (standard phosphate-buffered saline pH 7.4 + 0.1% triton x 100). Blocking and antibody incubation buffer is PBST +5% milk. Antibody dilutions varied from 1:1000 to 1:2000.

RESULTS

Three SULU1 antisera (against both 539A and 540A) and two SULU3 antisera (542A) reacted specifically with the peptide antigens. Antisera binding was competable with peptide. Experiments with extracts from cells transfected with epitope-tagged SULU1 and SULU3 genes are underway.

Antisera against the PAK4 C-terminal peptide 554A reacted with purified Gst-PAK4 and detected a protein of the correct molecular weight from tissue culture cells. Specific immunoprecipitation experiments are ongoing to determine the reactivity with native protein.

Similar immunization and antisera testing experiments are underway for each of the other novel STE20-kinases.

STE20-related protein kinase peptide immunogens and their specificity in recognizing endogenous protein by Western blots or immunoprecipitations.

Protein	Sequence	Aa positions	Conj	West.	IP
STLK2	EKFQKCSADESP	405-416	KLH	Y	Y
STLK4	SISNSELFPTTDVG	252-267	KLH	Y	Y
SULU1	LDFPKEDYR	890-898	KLH	Y	Y
SULU1	HGDPRPEPRPTQ	409-420	KLH	Y	Y

SULU3	PSTNRAGSLKDPEC	2-14	KLH	N	ND
SULU3	DPRTRASDPQSPPQVSRHK	411-429	KLH	ND	ND
PAK4	CLVPLIQLYRKQTSTC	666-680	KLH	ND	Y
PAK5	PLMRQRNRTR	390-398	KLH	Y	Y
PAK5	SGDERRRAGPEKRPKSS	148-163	KLH	Y	Y
PAK5	(C) RRKSLVGT PYWMAPE	471-485	KLH	Y	ND

ND=not done yet

STE20-related protein kinase GST fusion protein immunogens and their specificity in recognizing endogenous protein by Western blots or immunoprecipitations.

5

Protein	domain	Aa positions	West.	IP
ZC1	Coiled-coil/pro/B/C	350-867	Y	Y
ZC1	B	615-732	Y	Y
ZC2	Coiled-coil /pro/B	348-762	ND	ND
ZC2	B	658-762	Y	Y
PAK4	Nterm	252-426	ND	ND
PAK4	Kinase/Cterm	350-681	ND	Y
PAK5	A/ Nterm	53-330	ND	ND
PAK5	A/Nterm	53-309	ND	ND

ND=not done yet

The 50kD STLK2 protein was expressed highly in several hematopoietic cell lines including Jurkat, pGL10, Ramos, A20, WEHI-231, K562, HEL and freshly isolated thymocytes from C57/BL6 mice. High levels of STLK2 expression were also detected in several tumor cell lines including Calu6, Colo205, LS180, MDAM231 and A549.

10 The 160 kD ZC1 protein was detected in Jurkat T cells, Colo205, HCT116, RIE-1, 293T, MDAMB231, and SK-MEL28.

15 The 170 kD ZC2 protein was detected in SK-Mel28 and UACC-62.

20 Elevated levels of the 64 kD PAK5 protein were confirmed in the breast cancer cell lines MDA-231 and MCF-7, and in the lung cancer cell line A549.

160

Example 5: Recombinant Expression and Biological Assays
for STE20-related Protein Kinases.

Materials and Methods

Transient Expression of the Ste20-related Kinases in
Mammalian Cells

The pcDNA expression plasmids (10 µg DNA/100 mm plate) containing the STE20-related kinase constructs are introduced into 293 cells with lipofectamine (Gibco BRL). After 72 hours, the cells are harvested in 0.5 mL solubilization buffer (20 mM HEPES, pH 7.35, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1.5 mM MgCl₂, 1 mM EGTA, 2 mM phenylmethylsulfonyl fluoride, 1 µg/mL aprotinin). Sample aliquots were resolved by SDS polyacrylamide gel electrophoresis (PAGE) on 6% acrylamide/0.5% bis-acrylamide gels and electrophoretically transferred to nitrocellulose. Non-specific binding was blocked by preincubating blots in Blotto (phosphate buffered saline containing 5% w/v non-fat dried milk and 0.2% v/v nonidet P-40 (Sigma)), and recombinant protein was detected using the various anti-peptide or anti-GST-fusion specific antisera.

In Vitro Kinase Assays

Three days after transfection with the STE20-related kinase expression constructs, a 10 cm plate of 293 cells was washed with PBS and solubilized on ice with 2 mL PBSTD containing phosphatase inhibitors (10 mM NaHPO₄, pH 7.25, 150 mM NaCl, 1% Triton X-100, 0.5% deoxycholate, 0.1% SDS, 0.2% sodium azide, 1 mM NaF, 1 mM EGTA, 4 mM sodium orthovanadate, 1% aprotinin, 5 µg/mL leupeptin). Cell debris was removed by centrifugation (12000 x g, 15 min, 4 °C) and the lysate was precleared by two successive incubations with 50 µL of a 1:1 slurry of protein A sepharose for 1 hour each. One-half mL of the cleared

supernatant was reacted with 10 μ L of protein A purified kinase-specific antisera (generated from the GST fusion protein or antipeptide antisera) plus 50 μ L of a 1:1 slurry of protein A-sepharose for 2 hr at 4 °C. The beads were 5 then washed 2 times in PBSTDS, and 2 times in HNTG (20 mM HEPES, pH 7.5/150 mM NaCl, 0.1% Triton X-100, 10% glycerol).

The immunopurified kinases on sepharose beads were resuspended in 20 μ L HNTG plus 30 mM MgCl₂, 10 mM MnCl₂, and 10 20 μ Ci [α^{32} P]ATP (3000 Ci/mmol). The kinase reactions were run for 30 min at room temperature, and stopped by addition of HNTG supplemented with 50 mM EDTA. The samples were washed 6 times in HNTG, boiled 5 min in SDS sample buffer and analyzed by 6% SDS-PAGE followed by autoradiography. 15 Phosphoamino acid analysis was performed by standard 2D methods on ³²P-labeled bands excised from the SDS-PAGE gel.

Similar assays were performed on bacterially expressed GST-fusion constructs of the kinases.

20 ZC1 Assay buffer: 20 mM Tris pH 7.4, 200 mM NaCl, 0.5 mM DTT, 3 mM MgCl₂, 0.3 mM MnCl₂, 100 μ M ³²PyATP.

Substrates: myelin basic protein (MBP) at 0.28 mg/mL and phosphorylated ZC1 peptide RTVGRRNTFIGT-PPYWMAPE at 17 μ M (bold underlined residue shows site of phosphorylation).

25 At higher concentrations of MgCl₂ (3 mM), the activity of ZC1 (both full-length and recombinant kinase domain) is up to 10-fold greater towards exogenous substrate MBP. In contrast, the autophosphorylation and the phosphorylation of the activation loop peptide substrate are both inhibited. 30 Mn++ does not inhibit the autophosphorylation and the peptide phosphorylation by the truncated kinase domain form. However, both the MBP phosphorylation, Mn++-preferring activity AND the autophosphorylating, Mg++-preferring

activity are eliminated with mutation of the ATP-binding lysine in ZC1 (Lys54Ala) indicating that both activities are attributable to the ZC1 kinase domain.

5 SULU1 Assay buffer: This buffer is identical to that for ZC1, except for 5 mM MgCl₂. Under these conditions, other STE20 family members (PAK4, ZC1) were inhibited for autophosphorylation and required reducing the [Mn] to <0.3 mM for an efficient autophosphorylation reaction.

10 Substrates: MBP, phosvitin, or α -casein at 0.28 mg/mL.

PAK4, PAK5 Assay Buffer: 20mM Hepes pH 7.2, 130 mM KCl, 10 mM MgCl₂, 1 mM NaF, 20 mM B-glycerolphosphate, 0.5 mM DTT, 50 μ M ATP, 0.5 μ Ci ³²P γ ATP.

15 Substrates: MBP at 0.28 mg/mL and peptide substrates derived from PAK5 activation loop at 2.5 μ M.

20 STLK2 Assay buffer: Similar to that described above, except for the inclusion of 5 mM MgCl₂, 5 mM MnCl₂, and 5 μ Ci ³²P γ ATP.

Transformation (PAK experiments)

25 Low-passage NIH3T3 fibroblasts displaying normal morphology (flat, non-refractile cellular morphology), as well as low rates of spontaneous transformation, were used in transformation assays. NIH3T3 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% (v/v) fetal calf serum, penicillin (100 U/mL) and streptomycin (100 U/mL) and kept in an humidified incubator at 37 °C and 5% CO₂.

30 Cells were transfected with DNA-lipid complexes. As per manufacturer instructions, lipofectamine was utilized to transfect NIH3T3 cells. All transfections were with equal

5

amounts of plasmid DNA (DNA from the appropriate expression vector without insert was used to give equivalent amounts of DNA per transfection). 1 µg of activated allele of H-Ras was co-transfected with increasing amounts of various alleles of PAK5.

10

Foci were scored after 3 weeks by fixing 10 min in 10% methanol, 10% acetic acid for 10 min, followed by staining with 0.4% (w/v) crystal violet in 10% methanol for 10 min, and washing with deionized water and drying at room temperature.

Transfections, stimulations, and luciferase assays (ZC1 experiments)

15

Cells (10^7) were transiently transfected by electroporation using a Gene Pulser (Bio-Rad Labs) with the setting of 960 μ F and 250 V. 20-40 hours later, transfected cells (about 10^5) were stimulated with various stimuli. After a 6-hour stimulation, cells were lysed, and luciferase activities were measured using the MicroLumatPlus (EG&G Berthold). (J. Exp. Med. 183:611-620, 1996, hereby incorporated by reference herein in its entirety including any drawings, tables, or figures.)

20

25

RESULTS

Protein expression and kinase activity of novel STE20-related protein kinases

Protein	Observed size (kD)	Predicted Size (kD)	In vitro Kinase activity	Endogenous Kinase activity
STLK2	50	46	Y	Y
STLK4	55	50	Y	ND
ZC1	160	140	Y	Y
ZC2	170	150	Y	Y
KHS2	ND	101	ND	ND
SULU1	119	105	Y	Y
SULU3	140	115	ND	Y

184

PAK4	80	75	y	y
PAK5	64	64	y	y

ZC1: Regulation of kinase activity

5 ZC1 is constitutively active as a full-length kinase when expressed either *in vitro* (TNT rabbit reticulocyte system) or in NIH 3T3, 293T, or H1299 tissue culture cells. The endogenously expressed kinase is also active when immunoprecipitated from carcinoma cell lines.

10 ZC1 signaling Pathways

15 Using human leukemic T cell line Jurkat as a model system, the impact of cotransfected wild-type ZC1 on the activation of two reporter genes, RE/AP-luciferase and NF κ B luciferase, was examined. RE/AP is a composite in the IL-2 gene promoter containing both a NF κ B-like site and an AP-1 site.

20 Optimal activation of both RE/AP-luciferase and NF κ B-luciferase reporter genes in Jurkat T cells requires signals generated from stimulation of both T cell receptor and the costimulator receptor CD28. Cotransfection of wild-type ZC1 with either the RE/AP-luciferase or the NF κ B-luciferase reporter results in the activation of RE/AP or NF κ B when costimulated with the anti-T cell receptor monoclonal antibody or the pharmacological reagents PMA and ionomycin that bypass proximal T cell receptor. No activation was seen when costimulated with an anti-CD28 monoclonal antibody.

25 These results suggest that wild-type ZC1, when overexpressed, was replacing a CD28-specific signal to activate RE/AP or NF κ B. These results imply that ZC1 is involved in the CD28 signaling pathway. Since NF κ B is one of the major pathways also activated by the pro-inflammatory

cytokine TNF- α signaling, it is also likely that ZC1 may be a component in the TNF- α signaling pathways.

PAK5: Design of specific peptide substrates

5 To aid in the development of *in vitro* kinase assays for screening small molecule libraries to identify specific inhibitors, the search for specific peptide substrates for PAK5 was undertaken.

10 The rationale used to design such peptides is based on the hypothesis that upon binding activated small G protein, PAK5 undergoes a conformational change that results in derepression of its kinase activity followed by autophosphorylation on the activation loop resulting in a fully active kinase. The site of autophosphorylation for 15 related family members has been identified by biochemical and/or genetic means (e.g. Wu, C, et al. J.Biol.Chem 270:15984-15992 and Szczepanowska, et al. Proc.Natl.Acad.Sci 94, 8503-8508, 1997). Specific peptide substrates for PAK5 were designed from the sequence of the activation loop of 20 this kinase.

An activation loop PAK5 peptide phosphorylated on the Thr residue of the TPY motif served as a high-affinity substrate for PAK5.

25 PAK5 activation loop peptides as kinase substrates

Peptide #	Kinase	Sequence	Aa	SEQ ID	Kinase	substrate
1	PAK5	(C) RRKSLVGTPYWMAPE	471-485	102	PAK5	yes
2	PAK5	(C) RRK <u>SLVGTPYWMAPE</u>	471-485	102	PAK5	yes
3	PAK5	(C) RR <u>KSLVGTPYWMAPE</u>	471-485	102	PAK5	no
4	KHS1	KRKSFIGTPYWMAPE	171-185	U77129	PAK5	yes
5	STLK2	KRNTFVGT <u>PFWMAPE</u>	175-189	5	PAK5	poor
6	SULU1	PANSFVGTPYWMAPE	174-188	22	PAK5	poor
7	ZC1	RRNTF <u>IGTPYWMAPE</u>	184-198	13	PAK5	poor
8	ZC1	RRNTF <u>IGTPYWMAPE</u>	184-198	13	PAK5	poor
9	STLK4	RNKVRKTFVGTPCWM <u>APE</u>	66-83	7	PAK5	poor
10	PAK5	(C) RRK <u>SLVGTPYWMAPE</u>	471-485	102	PAK4	yes

Note: underlined/ bold residue was phosphorylated

Peptide #	Kinase	Notes
1	PAK5	Equally well as MBB
2	PAK5	High Km for PAK5 (1-10 μ M)
3	PAK5	S is the site of phosphorylation
4	KHS1	Similar to peptide 1
5	STLK2	
6	SULU1	
7	ZC1	
8	ZC1	Better than 7
9	STLK4	
10	PAK5	Same Km as phosph. by PAK5

PAK5: Transformation

5 Transformation of low-passage NIH3T3 cells by ras in the presence or absence of various alleles of PAK5 showed that the dominant negative, kinase-dead allele of PAK5 was able to block ras transformation of NIH3T3 cells. Thus, PAK5 activity is required for ras transformation of NIH3T3 cells.

10 Inhibition of PAK5 activity may have therapeutic value as an anti-proliferative agent for treating cancer.

PAK4 and PAK5: interaction with Cdc42

15 PAK 4 interacts with CDC42 small G-protein but not Rac, RhoA, or Ras as determined by co-transfection of recombinant genes and detection by kinase assays. PAK5 also interacts with Cdc42. Coding sequences of activated alleles of small G proteins (ras, Cdc42, Rac, Rho) tagged with a Myc epitope were transiently expressed in 293T cells, various alleles of 20 35S-labeled PAK5 tagged with HA epitope were expressed in vitro with the reticulocyte (TNT) system.

Example 6: Chromosomal Localization of Ste20-Related Protein Kinases

Materials And Methods

STE20 protein kinases STLK3, STLK4, ZC1, ZC2, ZC3, 5 KHS2, SULU1, PAK4, and PAK5 were mapped using the GeneBridge 4 Radiation Hybrid Panel, RH02.05 (Research Genetics). The GeneBridge 4 Panel consists of 91 hybrid panel samples, in addition to one human positive control (HFL), and one hamster negative control (A23). The standard reaction 10 conditions used to test and conduct PCR reactions using the GeneBridge 4 Panel are available from Research Genetics.

Oligonucleotide sequences (all 5' to 3') used for PCR mapping were:

15	STLK3: CTCCCATTCTAGCAAAATCA,	AGAGGCAGTATTGTCAGATGTA
	STLK4: CCACACATGCGTATCTCTGTTG,	TTGCTAGAACATCACATCAGGTACA
	ZC1: ATCCCTGGATCACACTGCTTCT,	CAAGGTGTTCTTGCCCTCTGTT
	ZC2: AGATGGACTGTACTGGGAGGG,	AGAAGAGCACTTGGCACTTATC
	ZC3: CATCATGAACCTGGTGACGGG,	CCAGTGAATCAAACCAGTAAAA
20	SULU1: CAAAACCTGCCGTCTCTTCTATT,	ATTGTGCTACTGGGATTCTGTG
	KHS2: GAATAGCGGTACCATGATAGAATA,	TACCAAAAGAGGCCAAAGTGTG
	PAK4: CTCAGTATTCTCTCCAAAGATTG,	GATGTTCTCTCCATTCTGTAAAG
	PAK5: CATCACTGGAAGTCTGCAGTG,	CAGGTGCAGTAGTCATTGC

25 Positive reactions were assigned a score of "1", negative reactions are assigned a score of "0", and ambiguous reactions are assigned a score of "2". Results were submitted to the Whitehead Institute (www.genome.wi.mit.edu) for position analysis. Chromosomal localizations for ZC4, SULU3, STLK2, STLK5 and STLK6 were 30 available publicly (for example, from Unigene). The chromosomal locations of GEK2 and STLK7 have not been determined.

STLK2_h	Xq25-27.1	(Public)
STLK3	2q31.3	(Sugen)
STLK4_h	3p22.3-p22.2	(Sugen)
STLK5_h	17q23.2-24.2	(Public)

<u>STLK6</u> h	2q32.2 -q33.3	(Public)
<u>STLK7</u> h	NA	
<u>ZC1</u> h	2p11.2	(Sugen)
<u>ZC2</u> h	3q26.31-3q26.32	(Sugen)
<u>ZC3</u> h	17p13.2-13.3	(Sugen)
<u>ZC4</u> h	Xq22	(Public)
<u>KHS2</u> h	2p22-2p22.2	(Sugen)
<u>SULU1</u> h	12q24.21	(Sugen)
<u>SULU3</u> h	17p11.2	(Public)
<u>GEK2</u> h	NA	
<u>PAK4</u> h	15q14	(Sugen)
<u>PAK5</u> h	19q13.2-q13.3	(Sugen)

Many of the STE 20 kinases were mapped to regions associated with various human cancers, as shown below.

The regions were also cross-checked with the Mendalian Inheritance in Man database, which tracks genetic information for many human diseases, including cancer.

References for association of the mapped sites with chromosomal abnormalities found in human cancer can be found in: Knuutila, et al., Am J Pathol, 1998, 152:1107-1123, hereby incorporated herein by reference in its entirety including any figures, tables, or drawings. Association of these mapped regions with other diseases is documented in the Online Mendalian Inheritance in Man (OMIM) (<http://www.ncbi.nlm.nih.gov/htbin-post/Omim>).

15

STLK2 h, Xq25-27.1, (Public)

Osteosarcoma, Xq25-qter, 2 of 31.

Lymphoproliferative syndrome, X-linked (OMIM No. 308240)

human STLK3, 2q31.3, (Sugen)

20

Squamous cell carcinoma of Head and Neck, 3 of 30.

STLK4 h, 3p22.3-p22.2, (Sugen)

Mantle cell lymphoma 3p14-p22 1 of 27

Squamous cell carcinoma of Head and Neck 3p22-p24 1 of 14

25

Cardiomyopathy, dilated (OMIM 601154)

STLK5_h, 17q23.2-24.2, (Public)

Cervical cancer, 17q, 1 of 30

Gastroesophageal junction adenocarcinoma xenograft, 17q, 1
5 of 5

Breast carcinoma, 17q12-qter, 1 of 16

Bladder carcinoma, 17q22-q23, 1 of 14

Breast carcinoma, 17q22-q25, 8 of 101

Non-small cell lung cancer, 17q24-q25, 6 of 50

10 Testis, 17q24-qter, 2 of 11

Malignant peripheral nerve sheath tumors, 17q24-qter, 5 of 7

Alzheimer disease, susceptibility to (OMIM 106180)

STLK6_h, 2q32.2 -q33.3, (Public)

15 Non-small cell lung cancer, 2q31-q32, 1 of 50

Squamous cell carcinoma of Head and Neck, 2q31-q33, 3 of 30

Small cell lung cancer, 2q32-q35, 1 of 22

ZC1_h, 2p11.2, (Sugen)

20 non-small cell lung cancer, 2pter-q13, 1 of 10

non-small cell lung cancer, 2pter-q21, 1 of 10

Pulmonary alveolar proteinosis, congenital (OMIM 178640).

ZC2_h, 3q26.31-3q26.32, (Sugen)

25 Non-small cell lung cancer, 3q26.1-q26.3, 26 of 103

Cervical cancer, 3q26.1-q27, 4 of 30

Small cell lung cancer, 3q26.3-qter, 3 of 35

Squamous cell carcinoma of Head and Neck, 3q26.3-qter, 3 of
13

30 Marginal zone B-cell lymphoma, 3q26-q27, 1 of 25

Parosteal osteosarcoma, 3q26-q28, 1 of 1

Gastrointestinal stromal tumor, 3q26-q29, 1 of 16

Mantle cell lymphoma, 3q26-q29, 1 of 5

ZC3_h 17p13.2-13.3 (Sugen)

Malignant fibrous histiocytoma of soft tissue, 17p, 2 of 58

Leiomyosarcoma, 17p, 7 of 29

5 Non-small cell lung cancer, 17p, 1 of 50

ZC4_h, Xq22, (Public)

Diffuse large cell lymphoma, Xq22-ter, 1 of 32

Deafness, X-linked 1, progressive. (OMIM 304700).

10

KHS2_h, 2p22-2p22.2, (Sugen)

Synovial sarcoma, 2p21-q14, 1_of_67

Follicular lymphoma, 2p22-p24, 1_of_46

15 Colorectal cancer, hereditary, nonpolyposis, type 1, Ovarian cancer (MSH2, COCA1, FCC1). (OMIM 120435).

SULU1_h, 12q24.21 (Sugen)

Neuroglial tumors, 12q22-qter, 1_of_15

20 Gastroesophageal junction adenocarcinoma, 12q23-qter, 1 of 5.

Non-small cell lung cancer, 12q24.1-24.3, 2 of 50.

SULU3_h 17p11.2 (Public)

Malignant fibrous histiocytoma of soft tissue, 17p, 2_of_58

25 Leiomyosarcoma, 17p, 7_of_29

non-small cell lung cancer, 17p, 1_of_50

Diffuse large cell lymphoma, 17p11.2, 1_of_32

Osteosarcoma, 17p11.2-p12, 4_of_31

30 PAK4_h: 15q14 (Sugen)

Schizophrenia, (OMIM 118511).

PAK5_h: 19q13.2-q13.3 (Sugen)
Follicular lymphoma, 19q13, 1 of 46*
Mantle cell lymphoma, 19q13, 1 of 5
Hepatocellular carcinoma, 19q13.1, 2 of 50
5 Small cell lung cancer, 19q13.1, 10 of 35
Breast carcinoma, 19q13.1-qter, 1 of 33
cervical cancer, 19q13.1-qter, 1 of 30
Testis, 19q13.1-qter, 1 of 11
Chondrosarcoma, 19q13.2, 1 of 29.
10 Malignant fibrous histiocytoma of soft tissue, 19q13.2-qter,
2 of 58
Non-small cell lung cancer, 19qcen-q13.3, 6 of 104

15 Example 7: Demonstration Of Gene Amplification By
Southern Blotting

Materials and Methods

Nylon membranes were purchased from Boehringer Mannheim. Denaturing solution contains 0.4 M NaOH and 0.6 M NaCl. Neutralization solution contains 0.5 M Tris-HCL, pH 7.5 and 1.5 M NaCl. Hybridization solution contains 50% formamide, 6X SSPE, 2.5X Denhardt's solution, 0.2 mg/mL denatured salmon DNA, 0.1 mg/mL yeast tRNA, and 0.2 % sodium dodecyl sulfate. Restriction enzymes were purchased from Boehringer Mannheim. Radiolabeled probes were prepared using the Prime-it II kit by Stratagene. The beta actin DNA fragment used for a probe template was purchased from Clontech.

Genomic DNA was isolated from 20 different tumor cell lines: MCF-7, MDA-MB-231, Calu-6, A549, HCT-15, HT-29, Colo 205, LS-180, DLD-1, HCT-116, PC3, CAPAN-2, MIA-PaCa-2, PANC-1, AsPC-1, BxPC-3, OVCAR-3, SKOV3, SW 626 and PA-1, and from two normal cell lines: human mammary epithelial cells and human umbilical vein endothelial cells.

192

A 10 µg aliquot of each genomic DNA sample was digested with EcoR I restriction enzyme and a separate 10 µg sample was digested with Hind III restriction enzyme. The restriction-digested DNA samples were loaded onto a 0.7% agarose gel and, following electrophoretic separation, the DNA was capillary-transferred to a nylon membrane by standard methods (Sambrook, J. et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory).

10

PAK5 Amplicon:

A 600 base pair fragment (EcoR I - Sac I) of the PAK5 gene was used as a template for a radiolabeled DNA probe which was hybridized to the blots at 42 °C for 48 hours in hybridization solution using standard methods (supra). The blots were exposed to a phosphorimager screen for 4 days, then scanned and analyzed using a Molecular Dynamics Storm 840 phosphorimager. The relative mass and gene copy number values of the PAK5 DNA fragments were calculated from the band density values obtained. The blots were re-hybridized with a radiolabeled probe copied from a fragment of human beta actin DNA and developed as above to confirm the sample mass loading equivalency.

RESULTS

The PAK5 gene was determined to exhibit 3-fold amplification compared to the normal DNA copy number in PANC-1 (pancreatic epithelioid carcinoma) and OVCAR-3 (ovarian adenocarcinoma) human cell lines, and approximately 2 times the normal copy number in the BxPC-3 (primary pancreatic adenocarcinoma) human cell line.

Similar Southern analyses can be performed for other STE20 kinases.

193

Example 8: Detection Of Protein-Protein Interaction
Through Phage Display

Materials And Methods

5 Phage display provides a method for isolating molecular interactions based on affinity for a desired bait. cDNA fragments cloned as fusions to phage coat proteins are displayed on the surface of the phage. Phage(s) interacting with a bait are enriched by affinity purification and the 10 insert DNA from individual clones is analyzed.

T7 Phage Display Libraries

All libraries were constructed in the T7Select1-1b vector (Novagen) according to the manufacturer's directions.

15 Bait Presentation

Protein domains to be used as baits were generated as C-terminal fusions to GST and expressed in *E. coli*. Peptides were chemically synthesized and biotinylated at the 20 N-terminus using a long chain spacer biotin reagent.

Selection

Aliquots of refreshed libraries (10^{10} - 10^{12} pfu) supplemented with PanMix and a cocktail of *E. coli* 25 inhibitors (Sigma P-8465) were incubated for 1-2 hrs at room temperature with the immobilized baits. Unbound phage was extensively washed (at least 4 times) with wash buffer.

After 3-4 rounds of selection, bound phage was eluted in 100 μ L of 1% SDS and plated on agarose plates to obtain single plaques.

30 Identification of insert DNAs

Individual plaques were picked into 25 μ L of 10 mM EDTA and the phage was disrupted by heating at 70 °C for 10 min.

2 μ L of the disrupted phage were added to 50 μ L PCR reaction mix. The insert DNA was amplified by 35 rounds of thermal cycling (94°C, 50sec; 50°C, 1min; 72°C, 1min).

5 Composition of Buffer

10x PanMix

5% Triton X100

10% non-fat dry milk (Carnation)

10 mM EGTA

10 250 mM NaF

250 μ g/mL Heparin (sigma)

250 μ g/mL sheared, boiled salmon sperm DNA (sigma)

0.05% Na azide

Prepared in PBS

15

Wash Buffer

PBS supplemented with:

0.5% NP-40

25 μ l g/mL heparin

20

PCR reaction mix

1.0 mL 10x PCR buffer (Perkin-Elmer, with 15 mM Mg)

0.2 mL each dNTPs (10 mM stock)

0.1 mL T7UP primer (15 pmol/ μ L) GGAGCTGTCGTATTCCAGTC

25 0.1 mL T7DN primer (15 pmol/ μ L) AACCCCTCAAGACCCGTTAG

0.2 mL 25 mM MgCl₂, or MgSO₄, to compensate for EDTA

Q.S. to 10 mL with distilled water

Add 1 unit of Taq polymerase per 50 μ L reaction

30 LIBRARY: T7 Select1-H441

RESULTS

Phage display baits and interactors

5

Bait	Domain	Aa	Patent SEQ ID	CDNA library	Interactor	Sequence Range & SEQ ID
SULU1	Coiled-coil2	752-898	22	H441	GEK2 cc dom (1)	677-820 SEQ #26
SULU3	Coiled-coil2	755-898	23	H441	SLK isoform	M83780

(1) SULU1 cc1 also interacted to a lesser extent with the coiled-coil domain of an SLK isoform.

10

The phage display data suggest potential interactions of SULU3 with SLK and SULU1 with GEK2 through their coiled-coil domains. Therefore two members of the SULU subfamily of STE20 kinases interact with two members of a separate STE20 family, the prototype being SLK.

15

These results suggest a specificity in the interaction, and imply that these STE20 kinases may interact with each other through homo- and hetero-dimerization. Alternatively SULU-related kinases could act immediately up- or down-stream of the SLK-related kinases in a signaling cascade.

20

One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The molecular complexes and the methods, procedures, treatments, molecules, specific compounds described herein are presently representative of preferred embodiments are exemplary and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention are defined by the scope of the claims.

25

It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to

the invention disclosed herein without departing from the scope and spirit of the invention.

5 All patents and publications mentioned in the specification are indicative of the levels of those skilled in the art to which the invention pertains.

The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed herein. Thus, for example, in each instance 10 herein any of the terms "comprising", "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the 15 use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed.

In particular, although some formulations described 20 herein have been identified by the excipients added to the formulations, the invention is meant to also cover the final formulation formed by the combination of these excipients. Specifically, the invention includes formulations in which one to all of the added excipients undergo a reaction during 25 formulation and are no longer present in the final formulation, or are present in modified forms.

In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby 30 described in terms of any individual member or subgroup of members of the Markush group. For example, if X is described as selected from the group consisting of bromine,

WO 99/53036

PCT/US99/08150

chlorine, and iodine, claims for X being bromine and claims for X being bromine and chlorine are fully described.

Other embodiments are within the following claims.

What is claimed is:

CLAIMS

5 1. An isolated, enriched, or purified nucleic acid molecule encoding a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5.

10

2. The nucleic acid molecule of claim 1, wherein said nucleic acid molecule comprises a nucleotide sequence that:

(a) encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, or SEQ ID NO:107;

15 (b) is the complement of the nucleotide sequence of
20 (a);

(c) hybridizes under highly stringent conditions to the nucleotide molecule of (a) and encodes a naturally occurring kinase polypeptide;

25 (d) encodes a kinase polypeptide having the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:103, or SEQ ID NO:107, except that it lacks one or more, but not all, of
30 the following segments of amino acid residues: 1-21, 22-274, or 275-416 of SEQ ID NO:5, 1-31, 32-308, 309-489 or 490-516 of SEQ ID NO:6, 1-178 or 179-414 of SEQ ID NO:7, 1-22, 23-289, 290-526, 527-640, 641-896, or 897-1239 of SEQ ID NO:13, 1-255, 256-442, 443-626, 627-954, or 955-1297 of SEQ ID

199

NO:14, 1-255, 256-476, 477-680, 681-983, or 984-1326 of SEQ
ID NO:15, 1-13, 14-273, 274-346, 347-534, or 535-894 of SEQ
ID NO:18, 1-21, 22-277, 278-427, 428-637, 638-751, or 752-
898 of SEQ ID NO:22, 1-66, 67-215, 216-425, 426-539, 540-
5 786, or 787-887 of SEQ ID NO:23, 1-25, 26-273, 274-422, 423-
632, or 633-748 of SEQ ID NO:24, 1-51, 52-224, 225-393, 394-
658, or 659-681 of SEQ ID NO:29, 1-25, 26-281, 282-430, 431-
640, 641-754, 755-901, or 902-1001 of SEQ ID NO:31, 1-10,
10 11-321, or 322-373 of SEQ ID NO:97, 1-57, 58-369, or 370-418
of SEQ ID NO:99, 1-52, 53-173, 174-307, 308-572, or 573-591
of SEQ ID NO:103, or 1-33, 34-294, 295-337, 338-472, 473-
724, or 725-968 of SEQ ID NO:107;

(e) is the complement of the nucleotide sequence of

(d);

15 (f) encodes a polypeptide having the amino acid
sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7,
SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ
ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID
NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:103, or SEQ ID
NO:107 from amino acid residues 1-21, 22-274, or 275-416 of
20 SEQ ID NO:5, 1-31, 32-308, 309-489 or 490-516 of SEQ ID
NO:6, 1-178 or 179-414 of SEQ ID NO:7, 23-289, 290-526, 527-
640, 641-896, or 897-1239 of SEQ ID NO:13, 1-255, 256-442,
443-626, 627-954, or 955-1297 of SEQ ID NO:14, 1-255, 256-
25 476, 477-680, 681-983, or 984-1326 of SEQ ID NO:15, 1-13,
14-273, 274-346, 347-534, or 535-894 of SEQ ID NO:18, 1-21,
22-277, 278-427, 428-637, 638-751, or 752-898 of SEQ ID
NO:22, 1-66, 67-215, 216-425, 426-539, 540-786, or 787-887
30 of SEQ ID NO:23, 1-25, 26-273, 274-422, 423-632, or 633-748
of SEQ ID NO:24, 1-51, 52-224, 225-393, 394-658, or 659-681
of SEQ ID NO:29, 1-25, 26-281, 282-430, 431-640, 641-754,
755-901, or 902-1001 of SEQ ID NO:31, 1-10, 11-321, or 322-
373 of SEQ ID NO:97, 1-57, 58-369, or 370-418 of SEQ ID

NO:99, 1-52, 53-173, 174-307, 308-572, or 573-591 of SEQ ID NO:103, or 34-294, 295-337, 338-472, 473-724, or 725-968 of SEQ ID NO:107;

(g) is the complement of the nucleotide sequence of

5 (f);

(h) encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, or SEQ ID NO:107, except that it lacks one or more, but not all, of the domains selected from the group consisting of a C-terminal domain, a catalytic domain, an N-terminal domain, a spacer region, a proline-rich region, a coiled-coil structure region, and a C-terminal tail; or

10 15 (i) is the complement of the nucleotide sequence of
(h).

3. The nucleic acid molecule of claim 1, further comprising a vector or promoter effective to initiate transcription in a host cell.

20 25 4. The nucleic acid molecule of claim 1, wherein said nucleic acid molecule is isolated, enriched, or purified from a mammal.

5. The nucleic acid molecule of claim 4, wherein said mammal is a human.

30 6. A nucleic acid probe for the detection of nucleic acid encoding a kinase polypeptide in a sample, wherein said polypeptide is selected from the group consisting of STLK2,

STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5.

5 7. The probe of claim 6, wherein said polypeptide is
a fragment of the protein encoded by the amino acid sequence
set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID
NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID
NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID
NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID
10 NO:103, or SEQ ID NO:107.

15 8. A recombinant cell comprising a nucleic acid
molecule encoding a kinase polypeptide selected from the
group consisting of STLK2, STLK3, STLK4, STLK5, STLK6,
STLK7, ZC1, ZC2, ZC3, KHS2, SULU1, SULU3, GEK2, PAK4, and
PAK5.

20 9. The cell of claim 8, wherein said polypeptide is a
fragment of the protein encoded by the amino acid sequence
set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID
NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID
NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID
NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID
NO:103, or SEQ ID NO:107.

25 10. An isolated, enriched, or purified kinase
polypeptide selected from the group consisting of STLK2,
STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2,
SULU1, SULU3, GEK2, PAK4, and PAK5.

30 11. The polypeptide of claim 10, wherein said
polypeptide is a fragment of the protein encoded by the
amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6,

SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107.

5

12. The polypeptide of claim 10, wherein said polypeptide comprises:

(a) the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107;

(b) the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, except that it lacks one or more, but not all, of the following segments of amino acid residues: 1-21, 22-274, or 275-416 of SEQ ID NO:5, 1-31, 32-308, 309-489 or 490-516 of SEQ ID NO:6, 1-178 or 179-414 of SEQ ID NO:7, 1-22, 23-289, 290-526, 527-640, 641-896, or 897-1239 of SEQ ID NO:13, 1-255, 256-442, 443-626, 627-954, or 955-1297 of SEQ ID NO:14, 1-255, 256-476, 477-680, 681-983, or 984-1326 of SEQ ID NO:15, 1-13, 14-273, 274-346, 347-534, or 535-894 of SEQ ID NO:18, 1-21, 22-277, 278-427, 428-637, 638-751, or 752-898 of SEQ ID NO:22, 1-66, 67-215, 216-425, 426-539, 540-786, or 787-887 of SEQ ID NO:23, 1-25, 26-273, 274-422, 423-632, or 633-748 of SEQ ID NO:24, 1-51, 52-224, 225-393, 394-658, or 659-681 of SEQ ID NO:29, 1-25, 26-281, 282-430, 431-640, 641-754, 755-901, or 902-1001 of SEQ ID NO:31, 1-10, 11-321, or 322-373 of SEQ ID NO:97, 1-57, 58-369, or 370-418 of SEQ

ID NO:99, 1-52, 53-173, 174-307, 308-572, or 573-591 of SEQ ID NO:103, 1-24, 25-289, 290-397, 398-628, 629-668, 669-872, or 873-1227 of SEQ ID NO:105, or 1-33, 34-294, 295-337, 338-472, 473-724, or 725-968 of SEQ ID NO:107;

5 (c) the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107 from 10 amino acid residues 1-21, 22-274, or 275-416 of SEQ ID NO:5, 1-31, 32-308, 309-489, or 490-516 of SEQ ID NO:6, 1-178 or 179-414 of SEQ ID NO:7, 23-289, 290-526, 527-640, 641-896, or 897-1239 of SEQ ID NO:13, 1-255, 256-442, 443-626, 627-954, or 955-1297 of SEQ ID NO:14, 1-255, 256-476, 477-680, 15 681-983, or 984-1326 of SEQ ID NO:15, 1-13, 14-273, 274-346, 347-534, or 535-894 of SEQ ID NO:18, 1-21, 22-277, 278-427, 428-637, 638-751, or 752-898 of SEQ ID NO:22, 1-66, 67-215, 216-425, 426-539, 540-786, or 787-887 of SEQ ID NO:23, 1-25, 26-273, 274-422, 423-632, or 633-748 of SEQ ID NO:24, 1-51, 20 52-224, 225-393, 394-658, or 659-681 of SEQ ID NO:29, 1-25, 26-281, 282-430, 431-640, 641-754, 755-901, or 902-1001 of SEQ ID NO:31, 1-10, 11-321, or 322-373 of SEQ ID NO:97, 1-57, 58-369, or 370-418 of SEQ ID NO:99, 1-52, 53-173, 174-307, 308-572, or 573-591 of SEQ ID NO:103, 1-24, 25-289, 25 290-397, 398-628, 629-668, 669-872, or 873-1227 of SEQ ID NO:105, or 34-294, 295-337, 338-472, 473-724, or 725-968 of SEQ ID NO:107; or

 (d) the amino acid sequence set forth in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, or SEQ ID NO:107, except that it

lacks one or more, but not all, of the domains selected from the group consisting of a C-terminal domain, a catalytic domain, an N-terminal domain, a spacer region, a proline-rich region, a coiled-coil structure region, and a C-terminal tail.

5

13. The kinase polypeptide of claim 10, wherein said polypeptide is isolated, purified, or enriched from a mammal.

10

14. The kinase polypeptide of claim 13, wherein said mammal is a human.

15

15. The kinase polypeptide of claim 10, wherein said polypeptide is a STLK2, STLK3, STLK4, STLK5, STLK6, STLK7 polypeptide.

20

16. The kinase polypeptide of claim 10, wherein said polypeptide is a ZC1, ZC2, ZC3, or ZC4 polypeptide.

25

17. The kinase polypeptide of claim 10, wherein said polypeptide is a KHS2 polypeptide.

25

18. The kinase polypeptide of claim 10, wherein said polypeptide is a SULU1 or SULU3 polypeptide.

19. The kinase polypeptide of claim 10, wherein said polypeptide is a GEK2 polypeptide.

30

20. The kinase polypeptide of claim 10, wherein said polypeptide is a PAK4 or PAK5 polypeptide.

21. An antibody or antibody fragment having specific binding affinity to a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5 or a kinase domain polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5.

10 22. A hybridoma which produces an antibody having specific binding affinity to a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5.

15 23. A method for identifying a substance that modulates kinase activity comprising the steps of:
(a) contacting a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5 with a test substance;
(b) measuring the activity of said polypeptide; and
(c) determining whether said substance modulates the activity of said polypeptide.

25 24. A method for identifying a substance that modulates kinase activity in a cell comprising the steps of:
(a) expressing a kinase polypeptide in a cell, wherein said polypeptide is selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5;
(b) adding a test substance to said cell; and

204

(c) monitoring a change in cell phenotype or the interaction between said polypeptide and a natural binding partner.

5 25. A method for treating a disease or disorder by administering to a patient in need of such treatment a substance that modulates the activity of a kinase selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, 10 PAK4, and PAK5.

15 26. The method of claim 25, wherein said disease or disorder is selected from the group consisting of immune-related diseases and disorders, organ transplantation, myocardial infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders, and cancer.

20 27. The method of claim 25, wherein said substance modulates kinase activity *in vitro*.

25 28. The method of claim 27, wherein said substance is a kinase inhibitor.

30 29. A method for detection of a kinase polypeptide in a sample as a diagnostic tool for a disease or disorder, wherein said method comprises:

35 (a) contacting said sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a kinase polypeptide selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5, said probe comprising the nucleic acid

sequence encoding said polypeptide, fragments thereof, or the complements of said sequences and fragments; and

(b) detecting the presence or amount of the probe:target region hybrid as an indication of said disease.

5

30. The method of claim 29, wherein said disease or disorder is selected from the group consisting of immune-related diseases and disorders, organ transplantation, myocardial infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders, and cancer.

10

31. A method for detection of a kinase polypeptide in a sample as a diagnostic tool for a disease or disorder, wherein said method comprises:

15

(a) comparing a nucleic acid target region encoding said kinase polypeptide in a sample, wherein said kinase polypeptide is selected from the group consisting of STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5, or one or more fragments thereof, with a control nucleic acid target region encoding said kinase polypeptide, or one or more fragments thereof; and

20

(b) detecting differences in sequence or amount between said target region and said control target region, as an indication of said disease or disorder.

25

32. The method of claim 31, wherein said disease or disorder is selected from the group consisting of immune-related diseases and disorders, organ transplantation, myocardial infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders, and cancer.

STE20_h - - - - - MAHLRGFANQHSRV - - - DPEELFTKLDRIGKGSSFGEVYKGIDNHTK 43
MST3_h - - - - - MAHSPVQSGLPGMQNLKADPEELFTKLEKIGKGSFGEVFKGIDNRTQ 47
STLK2_h - - - - - MAHSPVAVQVPGMQNNIAADPEELFTKLERIGKGSFGEVFKGIDNRTQ 47
STLK3_h TAAPAPAAPAPAQAAQAVGWPICRDAYELQEVIGSGATAVVQAAALCKPRQ 57

STE20_h EVVAIKITIDLEEADEIEDEIEDIQQEITVLSQCDSPYITRYFGSYLKSTKLWIIIMEYLGG 100
MST3_h KVVAIKIIDLLEEAEDEEDEIEDIQQEITVLSQCDSPYVTKYGGSYLKDTKLWIIIMEYLGG 104
STLK2_h QVVAIKIIDLLEEAEDEEDEIEDIQQEITVLSQCDSSYVTKYYGSYLNKGSKLWIIIMEYLGG 104
STLK3_h ERVAIKRINLEKCCQTSMDELLKEIQAMSOCSHPNVVWTYYT\$FVVKDELWLVWKILSG 114

STE20_h GSALDLKP[G]P-----[EETYIATILREILKGLDYLHSERKIHLDIKAANVLL 148
MST3_h GSALDLLEP[G]P-----[DETQIATILREILKGLDYLHSEKKIHRDIKAANVLL 152
STLK2_h GSALDLRAGP-----[FDEFQIATIMLKEILKGLDYLHSEKKIHRDIKAANVLL 152
STLK3_h GSMILDIIKYIVNRGEHKNGV[EEAIIATILKEVLEGLDYLHRNGQIHRDLKAGNILL 171
STLK4_h -----[KSGVLDXSTIATILREVLEGLEYLHKXGQIHRDVKAGNILX 41

STE20_h SE[QGDVKLADEFGV-A-----[GQLTDTQIKRNTFVGTPFWMAPEVTKQSA-YDFKADI 198
MST3_h SE[HGEVKLADEFGV-A-----[GQLTDQIKRNTFVGTPFWMAPEVIKQSA-YDSKADI 202
STLK2_h SE[QGDVKLADEFGV-A-----[GOLTDQIKRNTFVGTPFWMAPEV1QQSA-YDSKADI 202
STLK3_h GEDGSVQIADEFGVSAFLATGGDVTRNKV-RKTFVGTPCWMAPEVMEQVRGYDFKADM 227
STLK4_h GEDGSVQIADEFGVSAFLATGGDITRNKV-RKTFVGTPCWMAPEVMEQVRGYDFKADI 97

STE20_h	W S L G I T A I E L A K G E P P P N S D L H P M R W L F L I P K N S P P T L E G - - - - - Q H S K P F K E F 246
MST3_h	W S L G I T A I E L A R G E P P H S E L H P M K V L F L I P K N N P P T L E G - - - - - N Y S K P L K E F 250
STLK2_h	W S L G I T A I E L A K G E P P P N S D M H P M R V L F L I P K N N P P T L V G - - - - - D F T K S F K E F 250
STLK3_h	W S F G I T A I E L A T G A A P Y H K Y P P M K V L M L T L Q N D P P T L E T G V E D K E M M K K Y G K S F R K L 284
STLK4_h	W S F G I T A I E L A T G A A P Y H K Y P P M K V L M L T L Q N D P P S L E T G V Q D K E M L K K Y G K S F R K M 154
STE20_h	V E A C L N K D P R F R P T A K E L L K H K F I T R Y T K K T S F L T E L I D R Y K R - - - - - - - - - - - 289
MST3_h	V E A C L N K E P S F R P T A K E L L K H K F I I R N A K K T S Y L T E L I D R Y K R - - - - - - - - - - - 293
STLK2_h	I D A C L N K D P S F R P T A K E L L K H K F I V K N S K K T S Y L T E L I D R F K R - - - - - - - - - - - 293
STLK3_h	L S L C L Q K D P S K R P T A A E L L K C K F F Q K A K N R E Y L I E K L L T R T P D I A Q R A K K V R R V P G S 341
STLK4_h	I S L C L Q K D P E K R P T A A E L L R H K F F Q K A K N K E F L Q E K T L Q R A P T I S E R A K K V R R V P G S 211
STE20_h	W K S E G H G E E S S - S E D S D I D G E A E D G E Q G P I W - - - - - - - - - - - 319
MST3_h	W K A E Q S H D D S S - S E D S D A E T D G Q A S G G S D S G - - - - - - - - - - - 323
STLK2_h	W K A E G H S D D E S D S E G S D S E S T S R E N N T H P E W - - - - - - - - - - - 324
STLK3_h	S G H L H K T E D G D W E S D D E M D E K - S E E G K A A F S Q E K S R R V K E E - - - - - - - - - - - 382
STLK4_h	S G R L H K T E D G G W E S D D E F D E E - S E E G K A A I S Q L R S P R V K E S I S N S E L F P T T D P V G T 267
STE20_h	- - - - - T F P P T I R P S P H S K L H K G T A L H S S Q K P A E P V K 350
MST3_h	- - - - - D W I F T I R E K D P K N L E N G A L Q P S D L D R N K M K D 354
STLK2_h	- - - - - S F T T V R K K P D P K K V Q N G A E Q D L V Q T L S - - - - - 351
STLK3_h	- - - - - N P E I A V S A S T I P E Q I Q S L S V H D S Q G P P N A N E 413
STLK4_h	LL Q V P E Q I S A H L P Q P A G Q I A T Q P T Q V S L P P T A E P A K T A Q A L S S G S Q E T K I P - - - - - 320

Fig. 1C

Ste20_h	MAHLRG.FANQHSRV-----	DPEELFTKLDRI[G]SFG[E]VYKGIDNHT[K]	43
T19A5.2_ce	TTTSSDELPRQADDSSMKWDRYYIQKLDP[EVIFTKQERIGRGSFGEVYKGIDNRT[G]	57	
pak_sp	-----	LLYRNFVKIGQGASGD[VYSARQVGIN	26
STLK5_h	-----	-----	0

Ste20_h	EVVAIKI IDLEEADEDIEDITQQEITIVL S Q C D S P Y I T R Y F G S Y L K S T K L W I I M E Y L G G	100	
T19A5.2_ce	RVVAIKI IDLEQADEDIEDITQQEIQVLSQCDSQYVTKYFGSFLKGSKLWIIIMEYLG[G] 114		
pak_sp	LSVAIKM N IN QQPKKF - IVNEILVUMKSHHHKNIVNFIDTFFYKSELWWMVMEYMR[G]	82	
STLK5_h	-----	-----	0

Ste20_h	GSA LDLLKPGP[EE TYIATILREILKGLDYLHSERKIH[RDIAANVLLSEQGDV[KL[A	157
T19A5.2_ce	GSA LDLT[KS[GKLD[DES[HIAVILREILKGL[EYHSERKIH[RDIAANVLVSEHGDV[KVA	171
pak_sp	GSL TEVVVTNNNTL[SEGQIAAI[CKETLEGLQH[LHENGIVH[RDIAKSDN[IL[LSLQGD[DI[KLT	139
STLK5_h	--LICTHFM DGMNNELAIAYI[QGV[KL[KALDYI[HMGYVHRISVKAASHI[ISVDGKVYLS	55

Ste20_h	DFG VAGQLTD[QIKRNTFVGT[P[F-----WMAPEVIKQSA - -YDF[KADIWSL GIT[A	205
T19A5.2_ce	DFG VAGQLTEVKKRITFVGS[P[F-----WMAPELIKQSS - -YDY[KADIWSL GIT[A	219
pak_sp	DFG FCAQIDSNNMTKRTTMVGYTPY-----WMAPEVVTKE - -YGF[KVDVW[SLGIMA	187
STLK5_h	GLRSNLSMISAGQRQRRVVFHDFPKYSVKVL[PWLSP[EVLQONLQGYDAKSDIY[SVGITA	112

Ste20_h	IELAK[GEP[PNS[D[LH[PW[RV[L[IPK - NSPPPTLEG - Q-----	238
T19A5.2_ce	IELANG[GEP[PNS[D[LH[PW[RV[L[IPK - NPPPV[QGSQ-----	253
pak_sp	IMVEGE[GEP[PNS[D[LH[PW[RV[L[IPKISRP[EL-----	222
STLK5_h	CELAN[GHV[PFKDMPATQMLLEKLN - GTVPC[LD - TSTIPAEEELTMSPSRSVANSGLS	167

Ste20_h	YTKKTSFLTELIDRYKRWKSEGHGEESSESDIDGEAEDGEQGPIWTFPPPTIRPSSP	329
T19A5.2_ce	AKKNS-----ILVDLIERAAEYRLRTGVSSSDSDLDDEDSDGGGTGSKWDY-PTVRGPR	338
Pak_sp	AVPV-----SLIPLIKSIHHSGK	275
STLK5_h	DSLTTSPPRPSNGDSPSHPYHRTFSPHFFHVEQCLQRNPDARPSASTLNHSFFKQ	224
Ste20_h	HSKLHKGTALHSSQKPAEPVKRQPRSQCLSTLVRPV-----FGELKEKHKQSG--GS	379
T19A5.2_ce	VSAADDGTVRQRTDPRRAQVDRRSPSGPGGTIVRG-----SPQVAAVAQLR--NS	388
STLK5_h	-IKRRA-[SEALPELLRPVTITNF-EGSQSQDHSGIFGL-----	260
Ste20_h	VGA[EELENNAFSLAESCPGISDKLMVHLVERVQRFSHNRNHLISTR	426
T19A5.2_ce	XXALDQLRHVFRDVEDSCPGENELIEELMQRIAVPQVSQSDLDAIIRRLLTPPS	443
STLK5_h	VTN[EELEVDDWEF	274

Fig. 2B

ZC504.4_ce
NIK_m
ZC1_h
ZC2_h
ZC3_h

MSSGEDE --- IDLNSSLRDPAGIFELIEVVGNGTYGQVYKGRHVKTQ[LAIAKIMNIN] 58
MANDSPAKSLVDIDLSSLRDPAGIFELVEVVGNGTYGQVYKGRHVKTQ[LAIAKIMNIN] 69
MANDSPAKSLVDIDLSSLRDPAGIFELVEVVGNGTYGQVHEGRHVKTGQLAAIKVMDT 60
----- AFGEVYEGRHVKTGQLAAIKVMDTG 28
----- AFGEVYEGRHVKTGQLAAIKVMDT 28

ZC504.4_ce
NIK_m
ZC1_h
ZC2_h
ZC3_h

DEED[EIKLEINMLKKHSHHRNIATYYGAFIKKLPSSTGKH]DQLWLVMEFCGS[SITDLVK 116
DEEEEIT[EIKLEINMLKKYSHHRNIATYYGAFIKKSPGHD]DQLWLVMEFCGASITDLVK 117
DEEEEI[KLEINMLKKYSHHRNIATYYGAFIKKSPGHD]DQLWLVMEFCGAGSITDLVK 118
DEEEEI[KLEINMLKKYSHHRNIATYYGAFIKKNPPGMND]DQLWLVMEFCGAGSVITDLIK 84
DEEEEI[KQLEINMLKKYSHHRNIATYYGAFIKKSPPGND]DQLWLVMEFCGAGSSVITDLVK 84

ZC504.4_ce
NIK_m
ZC1_h
ZC2_h
ZC3_h

6/76

NTKGGS[KEEWIAIYICREILRGYHLHQSKVIHRDIKGQNVLLTDSAEVVKLVDFGVSAQL 176
NTKGNTLKEDWIAIYSREILGLAHLH[H]VIHRDIKGQNVLLTENAEVKLVDFGVSAQL 177
NTKGNTLKEDWIAIYSREILRGHLH[H]VIHRDIKGQNVLLTENAEVKLVDFGVSAQL 178
NTKGNTLKEEWIAIYICREILRG[SHLHQ]KVIHRDIKGQNVLLTENAEVKLVDFGVSAQL 144
NTKGNA[LKE]D[CLAYICREILGLAHLHA]KVIHRDIKGQNVLLTENAEVKLVDFGVSAQL 144

ZC504.4_ce
NIK_m
ZC1_h
ZC2_h
ZC3_h

DKTVGRRNTFIGTPYWWMAPEVIA[DESPEATYDSDLWSLGITALEM]AEGHPPPLCDMHP 236
DRTVGRRNTFIGTPYWWMAPEVIA[DENPDATYDYSRSDLWSC]GITATAIEMAE[GPPPLCDMHP 237
DRTVGRRNTFIGTPYWWMAPEVIA[DENPDATYDYSRSDLWSC]GITATAIEMAE[GAPPLCDMHP 238
DRTVGRRNTFIGTPYWWMAPEVIA[DENPDATYDFFKSDLWSLGITALEM]AE[GAPPLCDMHP 204
DRTVGRRNTFIGTPYWWMAPEVIA[DENPDATYDYSRSDLWSLGITALEM]AE[GAPPLCDMHP 204

Fig. 3A

IRHSIKEHIDRNRRVKKQDADYEYSGGSEDDDEPSPNNRDDSESS--SMIPMDNTLRKGFFQ	353
VRIQLKDHIDRTRKKGEEKDTEYEYSGSEEEEVPEQGEPESSIVNVPGESTLRRDFL	356
VRIQLKDHIDRTRKKGEEKDTEYEYSGSEEEEVPEQGEPESSIVNVPGESTLRRDFL	357
VRIQLKDHIDRTRKKGEEKDTEYEYSGSEEEEVPEQGEPESSIVNLPRESTLRRDFL	321
VRIQLKDHIDRTRKKGEEKDTEYEYSGSEEEEVPEQGEPESSIMNVPGESTLRRDFL	322

K[QESSRGFAEPGAQQLRRRLPQQPAPAPFQYQQSRYVEPR	VKLRAVSSRGAADGP	413
RLQQENKERSEALRRQQQLQQEQQL--	REQEYKRQLLAERQKRI-	398
RLQQENKERSEALRRQQQLQQEQQL--	REQEYKRQLLAERQKRI-	399
RLQQANKERSEALRRQQQLQQEQQL--	RENNEEHKROQLLAERQKRI-	361
RLQOENKSNSSEALLKOOQOQOQOQ	ROPEAHIKHLHORQRRI-	364

ZC504.4_ce
NIK_m
ZC1_h
ZC2_h
ZC3_h

ZZC504.4_ce
NIK_m
ZZC1_h
ZZC2_h
ZZC3_h

ZC504.4_ce
NIK_m
ZC1_h
ZC2_h
ZC3_h

ZC504.4_ce
NIK_m
ZC1_h
ZC2_h
ZC3_h

3B
Fig.

ZC504_4_ce	PAPQQSRKMSSEP[]LITHVKPED[DVLASELSKMGG-----	508
NIK_m	RVEREQYIRRQLEEEQRH[]ETLQQQLLQEQAQMLL-----	493
ZC1_h	RVEREQYIRRQLEEEQRH[]EVLQQQLLQEQAQMLL-----	-
ZC2_h	RAEH[E]HQEY[K]RKQLEEQ-----	-
ZC3_h	QAEREQEYIRRQLEEEQRQ[]EILQQQLLQEQA[]ILLEYKRKQLEEQ-----	469
ZC504_4_ce	[]HHNGRSREES-MSPPP[P]PREASSITDTIDVGELD[DNGAQAEWDDLKDI[M] 561	
NIK_m	[]HDHRRRP[]PHAQ[]Q[]PPQQQDRSKPSFHAPEPKPHYDPA[]DRAREVQWSHLASILKN 548	
ZC1_h	[]YLLSLQHDHRRRPHQ[]HSQQQ[]PPQQ[]PSSFH[]HAPEPKAHYEPA[]DRAREVE[D]FRKTNHS 579	
ZC2_h	[]YLVSLQHQR-[]-QEQR[P]VEKKPLYHYKEGMSPSEKPAWAKE[]VEERSRLNRQS 484	
ZC3_h	[]YLKSLQQQQQQQQLQK-[]QQQQQLLP[GDRKPLYHYGRGMNPADKPAWAREVE[]TRMNKQQ 528	
ZC504_4_ce	[]PNKPL[PPTPTDGENTLVSDVRRNGNG-----	595
NIK_m	[]NYSPVSRSHSFSDPSPKFAHHHL[]SQDP[]CPP[]SQRSE[]GLS[]QSSDSKSEV[P]EPTQK-----	599
ZC1_h	[]SPEAQSKQ-----	-
ZC2_h	[]SPAMPHKV-----	-
ZC3_h	[]NSPLAKSK-----	-
ZC504_4_ce	[]PGS-TGPE[]PIPQASP[]GPPGPLSQT[]P[]M[]QRPVEPQEGPH 574	
NIK_m	[]-----	-
ZC1_h	[]-----	-
ZC2_h	[]LVAVKSQGPALTASQSV-----HEQPTKG[]LSGFQEALNVT-----SHRVEMPRQ 575	
ZC3_h	[]KSLVAHRVPLKPYAAPVPRSQSLQDQPTRNLAFFPASHDPPAIPAPTATPSARGAVIRQ 634	

ZC504.4_ce - - - - -
NIK_m - - - - -
ZC1_h - - - - -
ZC2_h - - - - -
ZC3_h - - - - -

- - - - - NSGHGAYKKIPEIRPGIISL-DDDS 623
- - - - - AWSRSDSDEVPPRVPVRTTSRSPV[SRRD[SPLQGGQQQNS 639
- - - - - PQVPVRTTSRSPV[SRRD[SPLQ-GSGQQN 648
- - - - - NSDPTSENPPLPTRIEKFDRSSWLRQEEDI[P[KV[PQRTT[SISPA[LARKNSP[GNGSALGPR 635
- - - - - NSDPTSEGPGPSPNPP - - - AWVRPD-NEAPP[KV[PQRTT[SISPA[LARKNSP[GNGSALGPR 689

ZC504_4_ce	SDNE-----	E GNEPLMFKPIVRCPF SIFFWFLS --	651
WNJK_m	QAGQ-----	RNSTS SIEPRILLWERVEKLVPRPG --	667
ZC1_h	SQAG-----	QRNSTSIEPRILLWERVEKLVPRPG --	676
ZC2_h	LGSQ-----	PIRA\$NPDLRRT EPILE SPL QRTSSG	665
ZC3_h	ARPRSN SAWQIYLQRR AERGT P KPPGPPNASSNPDLRSDPG --	WERS --	742

ZC504.4_ce	QPNNGFQ-----NSDSR----	SSIQHSFSNRDREK	SFGYFGGAGAGGGTVNRP-----755
NIK_m	RSLKPAGEV-----DLTAL-----AKELRAVED	RPPPHKVTDYSSSEESGT	ED777
ZC1_h	RPLKPA-----DLTAL-----AKELRAVED	DYSSSEESGT	DDVEQ783
ZC2_h	RPSRPASYKKAI-----DLTAL-----KAMPRPKKA	SSSEESSESEESEEE	GES785
ZC3_h	RPADFV-----L-----L-----L-----L-----L-----L-----837	SEEVESSEDDEEEGEG	

Fig. 3D

ZC504.4_ce - - - - -
NIK_m [E G ADD S T SG P E D T R A A S S P N L S N G E T T E S V K T M I V H D D V E S E P A M T P] -- SKE GT L I V - - - - - 755
ZC1_h [E G A D E S T S G P E D T R A A S S L N L S N G E T T E S V K T M I V H D D V E S E P A M T P] -- SKE GT L I V - - - - - 837
ZC2_h [E T H D G T V A V S D I P R L I P T G A P G S N E Q Y N V G M V G T H G L E T S H A D S F S G S I S R E G T L M I R E T 845
ZC3_h [G P A E G S] -- - - - - R D T P G G R D G D D S V S T M V H D V E E I T G T Q P P] -- Y G G G I M V V Q R T 885

ZC504.4_ce - - - - -
NIK_m - - - - -
ZC1_h - - - - -
ZC2_h - - - - -
ZC3_h - - - - -
S G E K K R S G H S D S N G F A G H I N L P D L V Q Q S H S P A G T P T E G L G R V S T H S Q E M D S G T E Y G M G S S 905
P E E R N L H A D S N G Y T - - - N L P D V V Q P S H S P T E N S K G Q S P P S K D G S G D Y Q S R G L V K A P G - 941

ZC504.4_ce - - - - -
NIK_m S S S S F T P F I D P R I L L Q I S P S S G T T V T S V V G F S C D G L R P E A I R Q D P T - - - R K G S V V V N V N P T 901
ZC1_h S S S S F T P F I D P R I L L Q I S P S S G T T V T S V V G F S C D G M R P E A I R Q D P T - - - R K G S V V V N V N P T 907
ZC2_h T K A S S F T P F V D P R V Y Q T S P T D E D E E D E E S S A A A L F T G E L L R Q E Q A K L N E A R K I S V V V N V N P T 965
ZC3_h - K S S F T I M E V D L G I Y Q P G G S G D S I P I T A L V G G E G T R L D Q L Q Y D V - - - R K G S V V V N V N P T 994
ZC4_h - - - - - N V N P L 5

ZC504.4_ce - - - - -
NIK_m G T P A E N D A P E I R K Y K K F S G E I L C A A L W G V N L L I G T D S G L M L L D R S G Q G K V Y P L I S R R R F 832
ZC1_h N T R P Q S D T P E I R K Y K K R F N S E I L C A A L W G V N L L V G T E S G L M L L D R S G Q G K V Y P L I S R R R F 961
ZC2_h N T R P Q S D T P E I R K Y K K R F N S E I L C A A L W G V N L L V G T E S G L M L L D R S G Q G K V Y P L I N R R R F 967
ZC3_h N T R P H S D T P E I R K Y K K R F N S E I L C A A L W G V N L L V G T E N G L M L L D R S G Q G K V Y P L I N R R R F 1025
ZC4_h N T R A H S E I T P E I R K Y K K R F N S E I L C A A L W G V N L L V G T E N G L M L L D R S G Q G K V Y G L I G R R R F 1054
Y V S P A C K K P L I H M Y E K E E T S E I C C G S L W G V N L L I G T R S I N L Y L M D R S G K A D I T K L I R R R P F 65

ZC504.4_ce	DQNTVLE[GQNI][ATISGRKRRIRVYYLSWLRKILRQKILRTEGAGSANTTEKRNGWVNWGDD	--	889
NIK_m	QQMDVLEGLNVLTIS[GKD]LRRVYYLSWLRNKILHNDPEV	--	1013
ZC1_h	QQMDVLEGLNVLTIS[GKD]LRRVYYLSWLRNKILHNDPEV	--	1019
ZC2_h	QQMDVLEGLNVLTIS[GKD]LRRVYYLSWLRNKILHNDPEV	--	1077
ZC3_h	QQMDVLEGLNLIITIS[GKRNKLLR]VYYLSWLRNKILHNDPEV	--	1106
ZC4_h	ROLQVLEP[PLNLITIS[GHKNRRLRVYH][T]WL[RN]KILNNDPES	--	120

ZC504_4_ce	- - - - -	[QGAIEFKEIVR]YERIKFLVVG	[LESSIEIYAWAPKPYHKFMSSFKSFGSLSHVPLIV]	944
NIK_m	- - - - -	[EGCVHYKVVKYERIKFLVIALKSSVEVYAWAPKPYHKFMMAFKSFGELLHKPLLV]	[GELLHKPLLV]	1068
ZC1_h	- - - - -	[EGCVHYKVVKYERIKFLVIALKSSVEVYAWAPKPYHKFMMAFKSFGELVHKPLLV]	[GELVHKPLLV]	1074
ZC2_h	- - - - -	[EGCIHYKVVKYERIKFLVIALKNAVETIYAWAPKPYHKFMMAFKSFADLQHKPLLV]	[FADLQHKPLLV]	1132
ZC3_h	- - - - -	[MEGGHRYVVKYERIKFLVIALKSSVEVYAWAPKPYHKFMMAFKSFADLPHRPLLV]	[FADLPHRPLLV]	1161
ZC4_h	- - - - -	[KAIDKLTGCEHFSYLOHEETTYIAIALKSSIHLYAWAPKSFDESTAIKVEPTLDHKPVT]	[KAIDKLTGCEHFSYLOHEETTYIAIALKSSIHLYAWAPKSFDESTAIKVEPTLDHKPVT]	180

ZC504.4_ce	DLTVE <u>DNA</u> R <u>LKV</u> LYGSTGGFHAIDLD <u>PNSNGMQLL</u> DLTVE <u>E</u> GQRLKVIYGS <u>CAG</u> FHA <u>V</u> DVDS <u>G</u> SYDIYLP <u>T</u> HICSIK <u>P</u> HAI <u>I</u> ILPNT <u>DGMELL</u>	1004 1128
NIK_m	DLTVE <u>E</u> GQRLKVIYGS <u>CAG</u> FHA <u>V</u> DVDS <u>G</u> SYDIYLP <u>T</u> HICSIK <u>P</u> HAI <u>I</u> ILPNT <u>DGMELL</u> DLTVE <u>E</u> GQRLKVI <u>FGSHT</u> GFIYDVS <u>G</u> NSYDIYTP <u>SHI</u> QNI <u>T</u> PHAI <u>I</u> VLL <u>P</u> KTD <u>DGMEML</u>	1134 1192
ZC1_h	DLTVE <u>E</u> GQRLKVI <u>FGSHT</u> GFIYDVS <u>G</u> NSYDIYTP <u>SHI</u> QNI <u>T</u> PHAI <u>I</u> VLL <u>P</u> KTD <u>DGMEML</u>	1221
ZC2_h	DLTVE <u>E</u> GQRLKVI <u>FGSHT</u> GFIYDVS <u>G</u> NSYDIYTP <u>SHI</u> QNI <u>T</u> PHAI <u>I</u> VLL <u>P</u> KTD <u>DGMEML</u>	1221
ZC3_h	DLTVE <u>E</u> GQRLKVI <u>FGSHT</u> GFIYDVS <u>G</u> NSYDIYTP <u>SHI</u> QNI <u>T</u> PHAI <u>I</u> VLL <u>P</u> KTD <u>DGMEML</u>	1221
ZC4_h	DLTVE <u>E</u> GQRLKVI <u>FGSHT</u> GFIYDVS <u>G</u> NSYDIYTP <u>SHI</u> QNI <u>T</u> PHAI <u>I</u> VLL <u>P</u> KTD <u>DGMEML</u>	1221

3F
Fig.

ZC504_4_ce	[CYDNEGVYVNTYGRM]TKN[VVIQWGEMPS]SVAYISTGQIMGWGNKAIEIRSVDTGHLDGV	1064
NIK_m	[CYEDGVYVNTYGRITKDVVVLQWGEMPTSVAYIRSNTGMWGEKAIEIRSVETGHLDGV	1188
ZC1_h	[CYEDGVYVNTYGRITKDVVVLQWGEMPTSVAYIRSNTGMWGEKAIEIRSVETGHLDGV	1194
ZC2_h	[CYEDGVYVNTYGRITKDVVVLQWGEMPTSVAYIRSNTGMWGEKAIEIRSVETGHLDGV	1252
ZC3_h	[CYEDGVYVNTYGRITKDVVVLQWGEMPTSVAYICSNQIMGWGEKAIEIRSVETGHLDGV	1281
ZC4_h	[LTfNA[ALSWEANEQLFKK]TILEMW[KDIPSSIAFFECTDRTTGWQKAIEYRSLSQSRVLESE	300
		300
ZC504_4_ce	[FMHKKAQKLKFLCERNDKVFFSSAKGGSCQIYFMTLNKPGLTNW	1109
NIK_m	[FMHKRAQRLLKFLCGRNNDKVFSSVRSGGSQVYFMTLGRTSLLSW	1233
ZC1_h	[FMHKRAQRLLKFLCERNDKVFFASVRSGGSQVYFMTLGRTSLLSW	1239
ZC2_h	[FMHKRAQRLLKFLCERNDKVFFASVRSGGSQVYFMTLGRTSLLSW	1297
ZC3_h	[FMHKRAQRLLKFLCERNDKVFFASVRSGGSQVYFMTLGRTSLLSW	1326
ZC4_h	[LKRRSIKKLRFLCTRGDKEFTSLRNHHHSRVYFMTLGRKLEEQSNYDV	349

Fig. 3G

KHS1_h	5	LRPAADILRRNPQQDYELVQRVGSGTYGDVYKARNVHTGELAAVKIKLEPGDDFSLIQQ	64
KHS2_h	1	MNPQFDLSRNPQEDFELIQRIGSGTYGDVYKARNVNNTGELAAIKVIKLEPGEDFAVVQQ	60

KHS1_h	65	EIFMVKECKHCNIVAYFGSYLSREKLWICMEYCGGSLQDIYHVTGPLSELQIAAYVCRET	124
KHS2_h	61	EIIMMKDCKHPNIVAYFGSYLRRDKLWICMEFCGGSLQDIYHVTGPLSELQIAAYVSRET	120

KHS1_h	125	LQGLAYLHTKGKMHHRDIKGANILLTDHGDKVLADEFGVAAKITATIARKSFIGTPYWMAP	184
KHS2_h	121	LQGLYYLHSKGKMHHRDIKGANILLTDNGHVKLADEFGVSAQITATIARKSFIGTPYWMAP	180

KHS1_h	185	EVAAVEKNGGYNQLCDIWAVGITAIELGELQPPMFDLHPMRALFLMSKSNFQPPPQLKDKT	244
KHS2_h	181	EVAAYERKGGYNQLCDLWAVGITAIELAELQPPMFDLHPMRALFLMTKSNFQPPPQLKDKM	240

KHS1_h	245	KWSSTFHNFVKIALTKNPKKRPTAERLLTHTFVAQPGLSRALAVELLDKVNNPDNHAHYT	304
KHS2_h	241	KWSNSFHHFVKMALTKNPKKRPTAEKLLQHPFVTQH-LTRSLAIELLDKVNNPD-HSTYH	298

KHS1_h	305	EADDDEPHAIIRHTIRSTNRARAERTASEINFDKLQFEPPPLRKETEARDEMGLS---	361
KHS2_h	299	DFDDDDPEPLVAVPHRIHSTSRRNVRREEKTRSEITFFGQVKFDPPPLRKETEPHHELPDSDF	358

13/76

Fig. 4A

KHS1_h	362	- - - - - SDPNFMLQWNP - - - - - FV - - - - -	D 375
KHS2_h	359	LDSSEEIYYTARSNLDLQLEYGQGHQGGYFLGANAKSLKSVEEEELHQRGHVAHLEDDEGD 418	
* * * * *			
KHS1_h	376	GANTGKSTSKRAIPPPPLPPP KPRISSYPED-NFPDEEKASTIKHCP--DSESRAPQILRRQ 432	
KHS2_h	419	DDESKHSTLAKAKIPPPKPKSIFIPQEMHSTEDEDENQGTIKRCPMSSGSPAKPSQVPPRP 478	
* * * * *			
KHS1_h	433	SSPSGGPVAEETSSIGNGDISKL-MSENTEGSA - - - - - QAPQLPRKNDKRDFFPKPAIN 484	
KHS2_h	479	PPPR - - - LPPPHKPVVALGNMGMSMFQLMGERDGSICQQQNEHRGTONLSRK-EKKDVVKPISN 534	
* * * * *			
KHS1_h	485	GLPPTPKVLMGACFSKVFDGCPLKINCATSWIHPDTKDQYIIFGTEDGIYTLNLNELHEA 544	
KHS2_h	535	GLPPTPKVHMGA CFSSKVFN G CPLKIHCASSWINPDTRDQYLIFGAEGIYTLNLNELHET 594	
* * * * *			
KHS1_h	545	TMEQLFPRKCTWLYYVINNNTLMSLSEGKTFQLYSHNLIALFEHAK-KPGGLAAHIQTHRFPD 603	
KHS2_h	595	SMEQLFPRRCTWLYYVMNNCLLSIS-GKASQLYSHNLPGFLFDYARQMQLPVAIPAHKLKD 653	
* * * * *			
KHS1_h	604	RILPRKFALTTKIPDTKGCHKCCIVRNPYTGHKYLCGALQSGIVLLQWYEPMQKFMLIKH 663	
KHS2_h	654	RILPRKFSSVSAKIPETKWCQKCCCVVRNPYTGHKYLCGALQTSIVLLEWVEPMQKFMLIKH 713	

14/76

Fig. 4B

KHS1_h 664 FDFPLPSPLNVFEMLVPIPEQEYPMVCVAISKGTESNQVVQFETINLNNSASSWFTEIAGGS 723
 KHS2_h 714 IDFPPIPCLRMFEMLVVPEQEYPLVCVGVSRGDFNQVVRFETVNPNSTSSWFTE-SDT 771

*
 KHS1_h 724 QQLDSIHVTQLERDTVLVCLDKFVKIVNLOGKLKSSKKLASELSFDRIESVVCLQDSVL 783
 KHS2_h 772 PQTNVTHTQLERDTILVCLDCCIKIVNLRKSSRKLSSELTFDFQIESIVCLQDSVL 831

*
 KHS1_h 784 AFWKHGMQGKSFKSDEVTQEISDETRVFRLGSDRVVVLESRPTENPTAHSNLYILAGHE 843
 KHS2_h 832 AFWKHGNQGRSFRSNEVTQEISDSTRIFRLLGSDRVVVLESRPTDNPTANSNLYILAGHE 891

*
 KHS1_h 844 NSY
 KHS2_h 892 NSY
 846
 894

15/76

SU1_U_ce	MAPAVLQKPGVIKDPSIA	ALFSNKDPEQRYQDLREI	GHSFGA	YFAYDKKKNEQTVAIKK	60
SU11_h	- - - - -	MRKGVLKDPEIIDDLYK	GHGSFGA	YFATNAHTNEVVAIKK	54
SU13_m	- - - - -	MPSTNRAGSLKDPEDPE	GHGSFGA	YFARDVRTNEVVAIKK	58

SULU_ce	M F S G K Q A V E K W N D I L K E V S F L N T V Y H P H I V D Y K A C F L K D T T C W L V M E Y C I G S A A D I V D V 120
SULU1_h	M S Y S G K Q O T H E K W Q D I L K E V K F L R Q L K H P N T I E Y K G C Y L K E H T A W L V M E Y C I G S A S D L L E V 114
SULU3_m	M S Y S G K Q O S T E K W Q D I I K E V K F L Q R I K H P N S I E Y K G C Y L R E H T A W L V M E Y C I G S A S D L L E V 118

SULU_ce	R K G M R E V E I I A A I C S Q T L D A L R Y L H S L K R I H D I K A G N I L L S D H A I V K L A D F G S A S L V D P 180
SULU1_h	H K K P L Q E V E I I A A I T H G A L H G L A Y L H S H A L I H R D I K A G N I L L T E P G Q V K L A D F G S A S M A S P 174
SULU3_m	H K K P L Q E V E I I A A I T H G A L Q G L A Y L H S H T M I H R D I K A G N I L L T E P G Q V K L A D F G S A S M A S P 178

SULU_ce	AQTFIGTPFFMAPEVILAMDEGH	YTDRADI	PPPLFS	SNAMSALYHIA	240
SULU1_h	ANSFVGTPYWMAPENVILAMDEGQYDGV	DQ	KPPLFNMN	NAMSALYHIA	234
SULU3_m	ANSFVGTPYWMAPENVILAMDEGQYDGV	DQ	KPPLFNMN	NAMSALYHIA	238
SULU3_h	-	-	-	-	24

SU11_U_ce	QNDPPTLSPIDTSEQP[EWSLE[FVQFID[KCLRKPAEER[RM[SAE[ECFRHPFIQRSSRPSDTIQUE300
SU11_U_h	QNDSPTLQSN-----EWTD[SFREV[DYCLQKIPQERPTSAELL[RHD[FV
SU11_U_m	QNESPTLQSNMN-----ESPTLQSNMN-----EWSDYFRNFV
SU11_U_h	QNESPTLQSN-----ESPTLQSN-----EWSDYFRNFV
SU11_U_m	QNESPTLQSN-----ESPTLQSN-----EWSDYFRNFV

SULU_ce	L I Q R T K N M V L E D N F Q Y K K M R K L M Y L D E T E G K E G S E G N G A S D D L D F H G N E A N S I G R A G D S	360
SULU1_h	L I Q R T K D A V R E L D N L Q Y R K M K K I L F	313
SULU3_m	L I Q R T K D A V R E L D N L Q Y R K M K K L F	309
SULU3_h	L I Q R T K D A V R E L D N L Q Y R K M K K L F	103

Fig. 5A

SULU_ce ASSRASA\$LTsFRSMQSSGGAGLLVSTNTTGMNDNVHGSSSGNGSSSSSSARRRPPIPS 420
 SULU1_h QETRNGPLNES 324
 SULU3_m QEAHNGPAVEA 320
 SULU3_h QEAHNGPAVEA 114

SULU_ce QMISSTTS[GVGTMPSSHGSV[GASITAIAVNPTPS[SEPIPTS[Q[PTSKSESS-SILETAHD 479
 SULU1_h QEDDE[DSEHGTSLNREMDS[LGSNH[STIPSMSV[STGSQSSSVNSM[QEVMD[ESSELVMWHD 384
 SULU3_m QEEEQDHGVTGVSQNSVGSNQ[STIPSMSI[SASQSSSVNSL[PDA[DDKS-[ELDMMEGD 379
 SULU3_h QEEEQDHGVTGVSQNSVGSNQ[STIPSMSI[SASQSSSVNSL[PDV[DDKS-[ELDMMEGD 173

SULU_ce DPLDTSI-----RAPVKDLHM[PHRAVKERIATLQNHNKFATLRSQRII 521
 SULU1_h ESTINS[SVVH[KKDHVFTTRDEAGHGDPRPEPRPTQS[VQSQALHYRNTERFATIKSASLV 444
 SULU3_m HTVMSNSV[VIHLKPEEEENYQEEGDPRTRASSDPQSP[PPQVSRRHKSHYRNREHFATIRTA[SLV 439
 SULU3_h HTVMSNSV[VIHLKPEEEENYREEGDPRTRASSDPQSP[PPQVSRRHKSHYRNREHFATIRTA[SLV 233

SULU_ce NQEQQEYTKENNMYEQMSK[YKHL[RQAHHKE[LQQFFERCALDREQLRVKMDRE[EQLTYY 581
 SULU1_h TRQIH[EHEQENELREQMSGYKRMRRQHQKQLWTLQKEVETHANNS 504
 SULU3_m TRQMQEHEQDSSELREQMSGYKRMRRQHQKQLWTLRDKDLETQRNNF 499
 SULU3_h TRQMQEHEQDSSELREQMSGYKRMRRQHQKQLWTLRDKDLETCRNNF 293

SULU_ce SK[EKMVRCSQNNELDKRKKDIEDGEKMMKKTKNSQNQQQMKLYSAQQLKEYKYNKEAQK 641
 SULU1_h SIEL[EKLAKKQVAIIEKEAKVAADEKKFQQQILAQKKD[TTFFLESQKQYKICKEKIK 564
 SULU3_m AAEMEKLIIKKHQAAMEKEAKVMANEEKKFQHQHIQAAQKKELNSFLESQKREYKLRK[EQLK 559
 SULU3_h AAEMEKLIIKKHQAAMEKEAKVMSNEEKKFQHQHIQAAQKKELNSFLESQKREYKLRK[EQLK 353

SULU_ce	R	I	R	S	L	N	M	-P	R	S	T	Y	E	A	M	K	E	V	K	A	D	L	N	R	V	K	D	A	R	E	N	D	F	D	E	K	L	R	A	E	L	D	E	I	V	R	Y	R	R	Q	Q	L	S	N	L	H	700
SULU1_h	E	E	M	N	E	D	H	S	T	P	K	K	Q	E	R	I	S	K	K	E	N	L	Q	H	T	Q	A	Q	E	E	A	H	L	T	Q	Q	R	L	Y	D	K	N	C	R	F	K	R	K	I	M	I	K	R	H	624		
SULU3_m	E	E	L	N	E	N	Q	S	T	P	K	K	Q	E	W	L	S	K	Q	K	E	N	I	Q	H	F	Q	A	Q	E	E	A	N	L	R	R	Q	Y	L	E	E	C	R	F	K	R	M	L	G	R	H	619					
SULU3_h	E	E	L	N	E	N	Q	S	T	P	K	K	Q	E	W	L	S	K	Q	K	E	N	I	Q	H	F	Q	A	Q	E	E	A	N	L	R	R	Q	Y	L	E	E	C	R	F	K	R	M	L	G	R	H	413					

SULU_ce	S Q N E Y T Q R Q Q D E L R K [K H A M] Q S R Q Q P R D L K I Q E A Q I R K Q Y R Q V V K T Q T R Q F K L Y L T Q M V Q V	820
SULU_h	N Q L E Y N K R R E L H R K H V M G L R Q Q P K N L K A M E M Q I K K Q F Q D T C K V Q T K Q Y K A L K N H Q L E V	744
SULU3_m	N Q L E Y N K R R E L R R K H V M E V R Q Q P K S L K S K E L Q I K K Q F Q D T C K I Q T R Q Y K A L R N H L E T	739
SULU3_h	H N Q L E Y N K R R E L R R K H V M E V R Q Q P K S L K S K E L Q I K K Q F Q D T C K I Q T R Q Y K A L R N H L E T	533

SULU_ce YPKDEQKELTSRLKQDQMVKVALLASQYESQIKWQDQDKTVKLESWQEQRVLSEKLEK 880
 SULU_h TPKNEHKTILKTLKDEOTRKLAIALAEQYEQSINEMMASQALRDEAQEAECQALRQLQQ 804
 SULU3_m TPKNEHKAITPKSEHKAVLKRLIKETOTRKLAIALAEQYDHSTQALRDEAQEAECQV 748
 SULU3_h TPKSHEKAVLKRLIKETOTRKLAIALAEQYDHSTQALRDEAQEAECQVLKMQQLQQ 593

SULU_ce	EELIAYQK	KTRATTLEE	QIKKERTALE	EERIGHTRAM	LEQKIIER	EQMGER	RRLKKEQI	940
SULU1_h	EMELLNAYO	SKIKMQTAEQ	QHERELQKLE	QRVSLRAH	LEQKIEEE	ELAALQKE	SERIKNL	864
SULU3_h	ELLLNAYO	SKIKMQAEAQ	QHDRELLRE	QRVSLRALL	LEQKIEEE	ELALQNER	TERIRSSL	653

1876

Fig. 5C

SULU_ce RD[RH SQ[ER H R L E N H F V RT[G S T S R S S G G I A P G V G N S S S I Q M A M
SULU1_h LER[Q E R E I E T F D M E S L R M G F G N L V T L D F P K E D Y R
SULU3_h LER[Q A R E I E A F D S E S M R L G F S N M V L S N L S P E A F S H S Y P G A S G W S H N P T G G P H W G H P M G 713

982

898

713

SULU3_h G P P Q A W G H P M Q G G P Q P W G H P S G P M Q G V P R G S S M G V R N S P Q A L R R T A S G G R T E Q G M S R S T S 773

787

Fig. 5D

LOK_m	1	MAFANFRRILRLSTFEKRKSREYEHVRRDLDPNDVWEIVGELGDGAFGKVYKAKNKEETGA	60
GEK2_h	1	MAFANFRRILRLSTFEKRKSREYEHVRRDLDPNEVWEIVGELGDGAFGKVYKAKNKEETGA	60

LOK_m	61	AAAKVIETKSEEELDYIVEIEILATCDHPYIVKL LGAYYDGKLWIMIECPGGAVDA	120
GEK2_h	61	AAAKVIETKSEEELDYIVEIEILATCDHPYIVKL LGAYYHDGKLWIMIECPGGAVDA	120

LOK_m	121	IMLELDRGLTEPQIQVVCRCRQMLEALNFLHKGRIIHRLDKAGNVLMTLLEGDIRLADFGVSA	180
GEK2_h	121	IMLELDRGLTEPQIQVVCRCRQMLEALNFLHSKRIIHRLDKAGNVLMTLLEGDIRLADFGVSA	180

LOK_m	181	KNLKTLQKRDSFIGTPYWMAPEVVLCETMKDAPYDVKADIWSLGITLIEMAQIEPPHEL	240
GEK2_h	181	KNLKTLQKRDSFIGTPYWMAPEVVMCETMKDTPYDVKADIWSLGITLIEMAQIEPPHEL	240

LOK_m	241	NPMRVVLKIAKSDPPTLLTPSKWSVEFRDFFLKIALDKNPETRPSAAQLLQHPFVSRVTSN	300
GEK2_h	241	NPMRVVLKIAKSDPPTLLTPSKWSVEFRDFFLKIALDKNPETRPSAAQLLEHPFVSSITSN	300

LOK_m	301	KALRELVAEAKAEVMEEEIEDGREDEEEEADAVDAPPPLVNHTQDSANVTQPSLDNSKLLQD	360
GEK2_h	301	KALRELVAEAKAEVMEEEIEDGRDEGEEEEADAVDAASTLENHTQNSSEVSPPSLNADKPLEE	360

20/76

Fig. 6A

L0K_m	361	S - STPLPPSQPQE PVNGPCSQPSQDGPLQTTS PADGLSKNDNDLKVPVPLRKSRPLSMSDA	419
GEK2_h	361	SPSTPLAPSQSQDSVNEPCSQPSGDRSLQTTSPVVAPGNENGGLAVPVPLRKSRPVSMSDA	420

L0K_m	420	R I QM D E E K Q I P D Q D E N P S P A A S K S Q K A N Q S R P N S S A L E T L G G E A L T N G G L E L P S S V T P S H	479
GEK2_h	421	R I Q V A Q E K Q V A E Q G G D L S P A A N R S Q K A S Q S R P N S S A L E T L G G E K L A N G S L E P P A Q A A P G P	480

L0K_m	480	S K R A S D C S N L S T S E S M D Y G T S L S A D L S L N K E T G S L S L K G S K L H N K T L K R T R R F V V D G V E V	539
GEK2_h	481	S K R D S D C S S L C T S E S M D Y G T N L S T D L S L N K E M G S L S I K D P K L Y K K T L K R T R K F V V D G V E V	540

L0K_m	540	S I T T S K I I S E D E K K D E E M R F I L R R Q E L R L L Q K E E H R N Q T Q L S S K H E L Q L E Q M H K R F E Q	599
GEK2_h	541	S I T T S K I I S E D E K K D E E M R F I L R R Q E L R L L Q K E E H R N Q T Q L S N K H E L Q L E Q M H K R F E Q	600

L0K_m	600	E I N A K K K F Y D V E L E N L E R Q Q K Q Q V E K M E Q D H S V R R K E E A K R I R L E Q D R D Y A K F Q E Q L K Q M	659
GEK2_h	601	E I N A K K K F F D T E L E N L E R Q Q K Q Q V E K M E Q D H A V R R R E E A R R I R L E Q D R D Y T R F Q E Q L K L M	660

L0K_m	660	K K E V K S E V E K L P R Q Q R K E S M K Q K M E E H S Q K K Q R L D R D F V A K Q K E D L E L A M R K L T T E N R R E	719
GEK2_h	661	K K E V K N E V E K L P R Q Q R K E S M K Q K M E E H T Q K K Q L L D R D F V A K Q K E D L E L A M K R L T T D N R R E	720

21/76

Fig. 6B

LOK_m	720	ICDKERDCLSKKQELLR DREALWEMEEHQLQERHQ LVKQQLK DQYFLQRHDLLRKHEKE	779
GEK2_h	721	ICDKERECLMKKQELLR DREALWEMEEHQLQERHQ LVKQQLK DQYFLQRHELLRKHEKE	780
LOK_m	780	REQMQR YNQRMMEQLKVRQQQE KARLPKIQ RSDGETRMAMYYKKSLHINGAGSASEQREKI	839
GEK2_h	781	REQMQR YNQRMIEQLKVRQQQE KARLPKIQ RSEGKTRMANYYKKSLHINGGGSAAE QREKI	840
LOK_m	840	KQFSQQEEEKRQKAERLQQQQKHEHQMRDMVAQCESNMSELQQLQNEKCYLLVEHETQKLK	899
GEK2_h	841	KQFSQQEEEKRQKSERLQQQQKHENQMRDMLAQCESNMSELQQLQNEKCHLLVEHETQKLK	900
LOK_m	900	ALDESHNQS LIKE	911
GEK2_h	901	ALDESHNQNOLKE	912

22/76

SUBSTITUTE SHEET (RULE 26)

Fig. 6C

PAK1_h MSNNNGLDIQDKPPAPPMRNTSTWIGAGSKDAGTL[NHGS[KPLPPNPEEKKKDRFYRSIL - 59
 PAK65_h MEETQQKSHELELLSA - - - - -
 PAK3_m - M S D S L D N E E K P P A P P I R - - - - - M N S N N R D S S A L N H S S K P L P M A P E E K N K K A R L R S I F P G 54

PAK1_h PGDK[TNKKEKERPEISSL[PSDFEHTIHVGFDA[VTFGEFTGMPEQWARL[LQTSNITKS----- 115
 PAK65_h GTEKG[S]KKKEKERPEISSL[PSDFEHTIHVGFD[VTFGEFTGMPEQWARL[LQTSNITKL----- 96
 PAK3_m GGD[KTNKKKEKERPEISSL[PSDFEHTIHVGFD[VTFGEFTGIP[EQWARL[LQTSNITKL----- 110
 PAK4_h - - - MFRK[KK[KRPEISSL[PSDFEHTIHVGFD[VTFGEFTGIP[EQWARL[LQTSNITKL----- 110

PAK1_h ----- 115
 PAK65_h ----- 96
 PAK3_m ----- 110
 PAK4_h SRITRVQLQPMKTVVRGSAMPVDGYISGLNDIQKLSVISSNTLRGRSPTSRRAQSLGL 116

PAK1_h ----- 115
 PAK65_h ----- 96
 PAK3_m ----- 110
 PAK4_h LGDEHWATDPDMYLQSPQSERTDPHGLYLSCNGGTPAGHKQMPWPEPQSPPRVLPLNGLAAK 176

PAK1_h ----- 127
 PAK65_h ----- 108
 PAK3_m ----- 122
 PAK4_h AQSLGPAAEFQGASQRCLQLGACLQSSPPGASPPGTNRHGMKAAKHGSSEARPPQSCLVGS 236

PAK1_h	[LEFYNSKKTSNSQKYM	SFT	--DKSAEDYNSSSNALNVKA	VSETPAVPPVSEDED	DDDDAT185
PAK65_h	[LKFYDS--NTVKQKYL	SFT	--PPEKDGFPSGTPALNAKG	T[EA]PAVVT---EEE	DDDEET160
PAK3_m	[LKFYDS]KETVNNQKYM	SFT	--SGDKSAHGYIAAHQSNT	KT[GSEPPPLAPPVSEEE	E182
PAK4_h	ATGRPGGEGPSPKT	RESSLKRLFRSMFLSTAA	TAPPSSSKPKGPPQSKPN	SSFRPPQK296	
PAK5_h					-ASGAKLAAGRPFFNT14

PAK1_h	- - - P P V I A P R P E H T K S V Y T R S V I E P L P V T P T R D V A T S P I S P T E N N T T P P D A L L N T E K 241
PAK65_h	- - - A P P V I A P R P D H T K S I Y T R S V I D P V P A P V G D S H V D G A A K S L - - - - - D K 202
PAK3_m	D D N E P P P V I A P R P E H T K S I Y T R S V U V E S I A S P A A P N K E D I P P S A E N A N S T T L Y R - - - N T D R 239
PAK4_h	D N P P S L V A K A Q S L P S D Q P V G I F S P L T T S D T S S P Q K S L R T A P A T G Q L P G R S S P A - - - G S P R 353
PAK5_h	Y P R A D T D H P S R G A Q G E P H D V A P N G P S A G G L A I P Q S S S S R P T R A G A P S P G V L G P H A S 74

PAK1_h	Q K K P K M S D E E I L E K	[R S I V S V G D P K K Y T R F E K I]	276
PAK65_h	Q K K P K M T D E E I H E K	[L R T I V S I G D P K K Y T R Y E K I]	237
PAK3_m	Q R K K P S K M T D E E I L E K	[L R S I V S V G D P K K Y T R L E K I]	274
PAK4_h	T W H A Q I S T S N I L Y P Q D P T V A K G A L A G E D T G V V T H E Q F K A A L R M V V D Q G D P R L L L D S Y V K I]	413	
PAK5_h	E P O I A P P A C T P A A P A V P G P P G P R S P O R V S H E Q F R A A L Q L V V D P G D P R S Y L D N F I K I]	134	

PAK1_h	G Q G A S G T V Y T A M D V A T G Q E V A I K Q M N L Q Q Q P K K E L I N E I I L V M R E N K N P N I V N Y L D S Y L V	336
PAK65_h	G Q G A S G T V F T A T D V A L G Q E V A I K Q I N L Q K Q P K K E L I N E I I L V M K E L K N P N I V N F L D S Y L V	297
PAK3_m	G Q G A S G T V Y T A L D I A T G Q E V A I K Q M N L Q Q Q P K K E L I N E I I L V M R E N K N P N I V N Y L D S Y L V	334
PAK4_h	G E G S T G I V C L A R E K H S G R Q V A V K M M D L R K Q Q R R E L L F N E V V I M R D Y Q H E N V V E M Y K S Y L V	473
PAK5_h	G E G S T G I V C L A R E K H S G R Q V A V K M M D L R K Q Q R R E L L F N E V V I M R D Y Q H E N V V E M Y N S Y L V	194

24/76

Fig. 7B

PAK1_h GDELVVVMEYLAGGSLTDVVTETCMD[GQIAAVCRECLQALESLHSNQVIHRDIKSNDNIL 396
 PAK65_h GDEL[F]VVVMEYLAGRSLTDVVTETCMD[AQIAAVCRECLQALEFLHANQVIHRDIKSNDNIL 357
 PAK3_m GDELVVVMEYLAGGSLTDVVTETCMD[GQIAAVCRECLQALDFLHSNQVIHRDIKSNDNIL 394
 PAK4_h GEELVVVLM[EFL]QGGALTDIVSQVRLNNEQIA[ATVCEAVLQALYLAQGVIHRDIKSNDNIL 533
 PAK5_h GDELVVVMEFLEGGALTDIVTHTRM[NEEQIAAVCLAVLQALSVLHAQGVIHRDIKSNDNIL 254

PAK1_h LGMDGSVKLTDFGFCAQITPEQSKRSTMVGTPYWMAPENVTRKAYGPKVVDIWSLGIMAIIE 456
 PAK65_h LGMEGSVKLTDFGFCAQITPEQSKRSTMVGTPYWMAPENVTRKAYGPKVVDIWSLGIMAIIE 417
 PAK3_m LGMDGSVKLTDFGFCAQITPEQSKRSTMVGTPYWMAPENVTRKAYGPKVVDIWSLGIMAIIE 454
 PAK4_h LTLDGRVKLSDFGFCAQISKDVP[KRS]LVGTPYWMAPEVISRSL[YATEVDIWSLGIMVIE 593
 PAK5_h LTHDGRVKLSDFGFCAOVSKE[VPRRKSLVGTPYWMAPELISRLPYGPEDVYDIWSLGIMVIE 314

PAK1_h MIEGEPYYLNENPLRALYLIATNGTPELQNP[EKL]SAIFRDFLNRCLEMDVEKGSAKELL 516
 PAK65_h MVEGEPYYLNENPLRALYLIATNGTPELQNP[EKL]SPIFRDFLNRCLEMDVEKGSAKELL 477
 PAK3_m MVEGEPYYLNENPLRALYLIATNGTPELQNP[EKL]SAVEHDFLNRCLEMDVDRRGSAKELL 514
 PAK4_h MVDGEPYYFSDSPYQAMKRLRDSPP[KLN]SHKVSPPVLRD[FLERMLVRDPQERATAQELL 653
 PAK5_h MVDGEPYYFNEPP[KAMKWMIRDNLLP]PRLKLNLSKSPSILKGELDRILLVRDPAQRATAEELL 374

PAK1_h QHQFLKIAKPLSSLTPLIAAAKEATKNNH
 PAK65_h QHPFLKLA[KPLSSLTPLIMAAKEA[MKSNR]
 PAK3_m QHPFLKLA[KPLSSLTPLIIAAKEA[IKNSSR]
 PAK4_h DHPFLLQTLPECLVPLIQLYRKQTSTC
 PAK5_h KHPFLAKAG[GPPASSIVPLMRQNRTR

SEQ ID NO: 5 STLK2 human Nterm=1-21 kin=22-274

Cterm=275-416

MAHSPVAVQVPGMQNNIADPEELFTKLERIGKGSFGEVKGIDNRTQQVVAIKIIDLEEA
 EDEIEDIQQEITVLSQCDSSYVTYYGSYLGSKLWIIMEYLGGSALDLLRAGPDEFQ
 IATMLKEILKGLDYLHSEKKIHRDIKAANVLLSEQGDVKLADFGVAGQLTDTQIKRNTFV
 GTPFWMAPEVIQQSAYDSKADIWSLGITAIELAKGEPPNSDMHPMRVLFLIPKNNPPTLV
 GDFTKSFKEFIDACLNKDPFRPTAKELLKHKFIVKNSKKTSYLTELIDRFKRWKAEGHS
 DDESSEGSDSESTSRENNTHPEWSFTVRKKDPKKVQNGAEQDLVQLSCLSMIITPA
 FAELKQQDENNASRNQAIIELEKSIAVAEEAACPGITDKMVKKLIEKFQKCSADESP

SEQ ID NO: 6 STLK3 human Nterm=1-31 kin=32-308 Cterm=309-489

(insert=327-352) tail=490-516

TAAPAPAAPAAPAPAPAPAAQAVGWPICRDAYELQEVIDSGATAVVQAALCKPRQERV
 AIKRINLEKCQTSMDELLKEIQAMSQCSPNVVYTSFVVKDELWLVMKLLSGGSMLDI
 IKYIVNRGEHKNGVLEEAIIATILKEVLEGDLHRNGQIHRDLKAGNILLGEDGSVQIA
 DFGVSAFLATGGDVTRNKVRKTFVGTPCWAPEVMEQVRGYDFKADMWSFGITAIELATG
 AAPYHKYPPMKVLMQLQNDPPTLETGVEDKEMMKYGSFRKLLSLCLQKDPSKRPTAA
 ELLKCKFFQKAKNREYLIKEKLTRTPDIAQRACKVRRVPGSSGHLHKTEDGDWEWSDDEM
 DEKSEEGKAAFSQEKSRRVKEENPEIAVSASTIPEQIQSLSVHDSQGPPNANEDYREASS
 CAVNLVRLRNSRKELNDIRFEFTPGRDTADGVSQELFSAGLVGDGHVVIVAAANLQKIVD
 DPKALKTLTFKLASGCDGSEIPDEVKLIGFAQLSVS

SEQ ID NO: 7 STLK4 human Nterm=absent, kin=1-178, Ctail=179-414,
 insert1=198-222, insert2=253-293

KSGVLDXSTIATILREVLEGLEYLHKXGQIHRDVKAGNILXGEDGSVQIADFGVSAFLAT
 GGDIRNKVRKTFVGTPCWAPEVMEQVRGYDFKADIWSFGITAIELATGAAPYHKYPPM
 KVLMQLQNDPPTLETGVQDKEMLKKGKSF瑞KMSLCLQKDPEKRPTAAELLRHKFFQK
 AKNKEFLQEKTQRAPTISERAKVRRVPGSSGRLHKTEDGGWEWSDDFDEESEEKAA
 ISQLRSPRVKESISNSELPFTDPVGTLQVPEQISAHPQPAGQIATQPTQVSLPPTAE
 PAKTAQALSSGSGSQETKIPISLVLRLRNSKKELENDIRFEFTPGRDTAEGVSQELISAGL
 VDGRDLVIVAANLQKIVEEPQSNSVTFKLASGVEGSDIPDDGKLIGFAQLSIS

SEQ ID NO: 8 STLK5 human Nterm=absent, kin=1-222(lacks N-term),
 Ctail=224-274

LICTHFMDGMNELAIAYILQGVLKALDYIHHMGYVHRSVKASHILISVDGKVYLSGLRSN
 LSMISHGQRQRVVHDFFPKYSVKVLPWLSPVELQQNLQGYDAKSDIYSGVITACELANGHV
 PFKDMPATQMLLEKLNGTVPCLLDTSTIPAEELTMSPSRVANSGLSDSLTTSPRPSNG
 DSPSHPYHRTFSFHFVEQCLQRNPDARPSASTLLNHSFFQIKRRASEALPELLRPV

Fig. 8A
26/76

TPITNFEGSQSQDHSGIFGLVTNLEELLEVDDWEF

SEQ ID NO: 13 ZC1 human 1/5/98 Nterm=1-22 kin=23-289
 coiled-coil=290-526 pro=527-640 B=641-896 Rab/Rac-BD=897-1239
 MANDSPAksLVDidLSSLRDPAGIFELVEVVGNGTyGQVYKGRHVKTGQLAAIKVMDVTE
 DEEEEEEIKLEINMLKKYSHHRNIATYYGAFIKKSPPGHDDQLWLVMЕFCAGSITDLVKNT
 KGNTLKEDWIAYISREILRGLAHLHIHHVIHRDIKGQNVLLENNAEVKLVDGFVSAQLDR
 TVGRRNTFIGTPYWMAPEVIACDENPDATYDYSRSDLWSCGITAIEMAEGAPPLCDMHPMR
 ALFLIPRNPPRLLKSKWSKKFFSFIEGCLVKNYMQRPSTEQLLKHPFIRDQPNERQVRI
 QLKDHIDRTRKKRGEKDETEYEYSGSEEEEEVPEQEGERPSSIVNVPGESTLRRDFRLRLQ
 QENKERSEALRRQQQLQEQQLREQEYKRQLLAERQKRIEQQQKEQRRRLEEQQRREAR
 RQQEREQRRREQEEKRRLEELERRKEEEERRRAEEEKRRVEREQEYIRRQLEEEQRHLE
 VLQQQLLQEQQAMLLCRWREMEEHRQAERLQRQLQQEYALLSLQHDHRRPHPHQHSQQPP
 PPQQERSKPSFHapePKAHYEPADRAREVEDRFRKTNHSSPEAQSKQTGRVLEPPVPSRS
 ESFSNGNSESVHPALQRPAEPQVPVRTTSRSPVLSRRDPLQGSGQQNSQAGQRNSTSIE
 PRLLWERVEKLVPRPGSGSSGSSNSGSQPGSHPGSQSGSGERFRVRSSSKSEGSPSQRL
 ENAVKKPEDKKEVFRPLKPADLTALAKELRAVEDVRPPHKVTDYSSSEESGTTDEEDDD
 VEQEGADESTSGPEDTRAASSLNLSNGETESVKTMIVHDDVESEPAAMPSKEGTTLIVRRT
 QSASSTLQKHKSSEFTPFIDPRLQISPSSGTTVTSVVGFSQCDGMRPEAIRQDPTRKGS
 VVNVPNTNRPQSDTPEIRKYKKRFNSEILCAALWGVNLLVGTESGLMLLDRSGQGKVYP
 LINRRRFQQMDVLEGNLVLTISGKKDKLRVYYLSWLRNKILHNDPEVEKKQGWTTVGDL
 EGCVHYKVVKYERIKFLVIALKSSVEVYAWAPKPYHKFMFKSFGEVHKPLLVDLTVEE
 GQLKVIYGSCAGFHADVDGSGSYDIYLPTHIQCISIKPHAIILPNTDGMELLVCYEDE
 GVYVNTYGRITKDVLQWGEMPTSVAIYIRSNQTMGWGEKAIEIRSvetGHLDGVFMHKRA
 QRLKFLCERNDKVFFASVRSGGSSQVYFMTLGRTLSLLSW

SEQ ID NO: 14 ZC2 human Nterm=missing kin=1-255 coiled-coil=256-442
 pro=443-626 B=627-954 Rab/RacBD=955-1297
 AFGEVYEGRHVKTGQLAAIKVMDVTGDEEEEEEIKQEINMLKKYSHHRNIATYYGAFIKKNP
 PGMDQLWLVMЕFCAGSVTDLIKNTKGNTLKEEWIAYICREILRGLSHLHQHKVIHRDI
 KGQNVLLENNAEVKLVDGFVSAQLDRTVGRRNTFIGTPYWMAPEVIACDENPDATYDFKS
 DLWSLGITAIEMAEGAPPLCDMHPMRALFLIPRNPAKRLKSKWSKKFQSFIESTCLVKNH
 SQRPAATEQLMKHPFIRDQPNERQVRIQLKDHDRTKKRGEKDETEYEYSGSEEEEEEND
 SGEPSSILNLPRESTLRRDFRLQLANKERSEALRRQQLEQQQRENEEHKRQLLAERQKR
 IEEQKEQRRRLEEQQRREKELRKQQEREQRRHYEEQMRREEERRRAEHEQEYKRKQLEEQ
 RQAERLQRQLKQERDYLVSLQHQRQEQRPVEKKPLYHYKEGMSPEKPAWAKEVEERSRL
 NRQSSPAMPHKVANRISDPNLPPRSESFSISGVQPARTPPMLRPVDPQIPLVAVKSQGP
 ALTASQSVHEQPTKGLSGFQEALNVTSHRVEMPRQNSDPTSENPLPTRIEKFDRSSWLR

Fig. 8B
 27/76

QEEDIPPKVPQRTTSISPALARKNSPGNGSALGPRLGSQPIRASNPDLRRTEPILESPHQ
 RTSSGSSSSSTPSSQPSSQGGSQPGSQAGSSERTRVRANSKSEGPVLPHPEAKVKPEE
 SRDITRPSRPASYKKAIDEDLTALAKELRELRIEETNRPMKVTDYSSSESESSEEEE
 EDGESETHDGTAVSDIPRLIPTGAPGSNEQYNVGVMVGHLETSHADSFGSISREGTL
 MIRETSGEKKRSGHSDSNGFAGHINLPDVQQSHSPAGTPTEGLGRVSTHSQEMDSGTEY
 GMGSSTKASFTPVDPRVYQTSPTEDEEDEESSAAALFTGELLRQEQAQLNEARKISVV
 NVNPTNIRPHSDTPEIRKYKKRFNSEILCAALWGVNLLVGTENGLMLLDRGQGKVYNLI
 NRRRFQQMDVLEGNLVLTISGKKNKLRYYLWLRLRNRLHNDPEVEKKQGWITVGDLEG
 CIHYKVVKYERIKFLVIALKNAVEIYAWAPKPYHKFMFKSFADLQHKPLVDLTVEEGQ
 RLKVIFGSHTGFHVIDVDSGNSYDIYTPSHIQGNITPHAIIVILPKTDGMEMLVCYEDEGV
 YVNTYGRITKDVLQWGEEMPTSVAIHSNQIMGWGEKAIEIRSVEVGHLGVFMHKRAQR
 LKFLCERNDKVFFASVRSGGSSQVFFMTLNRSMMNW

SEQ ID NO: 15 ZC3 human kin=1-255 coiled-coil=256-476 pro=477-680

B=681-983 Rab/RacBD =984-1326

AFGEVYEGRHVKTGQLAAIKVMDVTEDEEEIKQEINMLKKYSHHRNIATYYGAFIKKSP
 PGNDDQLWLVMFCGAGSVTDLVKNTKGNALKECDIAYICREILRGLAHLHAHKVIHRDI
 KGQNVLLTENAEVKLVDFGVSAQLDRTVGRRNTFIGTPYWMAPEVIACDENPDATYDYRS
 DIWSLGITAIEMAEGAPPLCDMHPMRALFLIPRNPPRLKSKKWSKKFIDFIDTCLIKTY
 LSRPPTEQLLKFPFIRDQPTERQVRIQLKDHDIDRSRKKRGEKEETEYEYSGSEEEDDSHG
 EEEGPSSIMNVPGESTLRREFLRLQQENKSNSEALKQQQLQQQQQRDPEAHIKHLLHQR
 QRRIEEQKEERRVEEQQREREQRKLQEKEQQRRLEDMQALRREEERRQAEREQEYIRH
 RLEEEQRQLEILQQQLLQEQQALLLEYKRKQLEEQRQSERLQRQLQQEHAYLKSLQQQQQQ
 QQLQKQQQQQLLPGDRKPLYHYGRGMNPADKPAWAREVEERTMNKQQNSPLAKSKPGST
 GPEPPPIPQASPGPPGPLSQTPPMQRPVEPQEGPHKSLVAHRVPLKPYAAPVPRSQSLQDQ
 PTRNLAAPASHDPDAIPAPTAUTPSARGAVIRQNSDPTSEGPGSPNPPAWVRPDNEAP
 PKVPQRTSSIATALNTGAGGSRPAQAVRARPRNSAWQIYLQRAERGTPKPPGPPAQ
 PGPPNASSNPDLRRSDPGWERSDSVLPASHGHLHQAGSLERNRVGVSSKPDSSPVLSGN
 KAKPDDHRSRPGRPADFVLLKERTLDEAPRPPKKAMDYSSSEEVESSEDDEEGEGGPA
 EGSRDTPGGRDGDTDSVSTMVHDVEEITGTQPPYGGTMVVQRTPEEERNLLHADSNGY
 TNLPDVVQPSHSPTENSKGQSPPSKDGSGDYQSRLVKA PGKSSFTMFVDLGIYQPGGSG
 DSIPITALVGGEGRQLDQLQYDVRKGSVVNPNTRAHSETPEIRKYKKRFNSEILCAA
 LWGVNLLVGTENGLMLDRSGQGKVYGLIGRRRFQQMDVLEGNLNLITISGKRNLRVYY
 LSWLRNKILHNDPEVEKKQGWTTVGDMECGHRYVVKYERIKFLVIALKSSVEVYAWAPK
 PYHKFMFKSFADLPHRPLLVDLTVEEGQRLKVIYGSSAGFHADVDGNSYDIYIPVHI
 QSQITPHAIIFLPNTDGMEMLLCYEDEGVYVNTYGRIIKDVVLQWGEEMPTSVAIICSNQI
 MGWGEKAIEIRSVEVGHLGVFMHKRAQRKLFLCERNDKVFFASVRSGGSSQVYFMTLNR
 NRIMNW

Fig. 8C

28/76

SEQ ID NO: 16 ZC4 human Nterm kin coiled-coil pro B=missing
Rab/RacBD=1-349

NVNPLYVSPACKPLIHMYEKEFTSEICCGSLWGVNLLGTRSNLYLMDRSGKADITKLI
RRRPFRQIQVLEPLNLLITISGHKNRLRVYHTLWRNKILNNDPESKRQEEMLKTEEAC
KAIDKLTGCEHFSQLQHEETTYIAIALKSSIHLYAWAPKSFDESTAIKVFP TL DHKP VTV
DLAIGSEKRLKIFFSSADGYHLIDA ESEVMSDVTLPKNPLEIIIIPQNIIILPDCLGIGMM
LTFNAEALSVEANEQLFKKILEMWKDIPSSIAFFECTQRTTGWGQKAIEVRSLQSRVLESE
LKRRSIKKLRFLTRGDKLFFTSTLRNHHSRVYFMTLGKLEELQS NYDV

SEQ ID NO: 18 KHS2 human Nterm=1-13 kin=14-273 A=274-346

Pro=347-534 RabBD =535-894

MNPGFDLSRRNPQEDFELIQRIGSGTYGDVYKARNVNTGELAAIKVIKLEPGEDFAVVQQ
EIIMMKDCCKHPNIVAYFGSYLRRDKLWICMEFCGGGSLQDIYHVTGPLSELQIAV SRET
LQGLYYLHSKGKMH RDIKGANILLTDNGHVKLADFGVSAQITATIAKRKSFIGTPYWMAP
EVAVERKGGYNQLCDLWA VGITAIELAELQPPMFDLHPMRALFLMTKSNFQPPKLKD KM
KWSNSFHHFVKMALT KNPKKRPTAEKLLQHPFVTQHLTRSLAIELL DKVNNPDHSTYHDF
DDDDPEPLVAVPHRIHSTS RNVREEKTRSEITFGQVKFDPLRKETEPHHELPDS DGFL D
SSEEIYYTARSNLDLQLEYGQGHQGGYFLGANKSLLKSVEELHQRGHVAHLEDDEGDDD
ESKHSTLAKA KIPPLPPKPKSIFIPQEMHSTEDENQGTIKRCPMMSGPAKPSQVPPR PPP
PRLPPHKPVALGN GMSSFQLNGERD GSLCQQQNEHRGTNLSRKEKKDVPKPISNGLPPT P
KVHMGACFSKVFN G CPLKIHCASSWINPDTRDQYLIFGAEGIYTLNLNELHETSMEQL F
PRRCTWL YVMNNCLLSISGKASQLYSHNLPGLFDYARQMQLPVAIPAHKL PDRILPRKF
SVSAKIPETKWCQKCCV VRNPYTGH KYLC GALQTSIVLLEWVEPMQKFMLIKHIDFPIPC
PLRMFEMLVVPEQEYPLVCVGVS RGRDFNQVVFETVNPNSTSSWFTESDTPQTNVTHVT
QLERDTILVCL DCCI KIVNLQGRLKSSRKLSSELTDFQIESIVCLQDSVLA FWK HGMQ G
RSFRSNEVTQEISDSTRIFRLLGSDRVVVLESRPTDNPTANSNLYILAGHENS Y

SEQ ID NO: 22 SULU1 human N=1-21 kin=22-277 A=278-427

coiled-coil1=428-637 B=638-751 coiled-coil2=752-898

MRKGVLKDPEIDDLFYKDDPEELFIGLHEIGHGSFGAVYFATNAHTNEVVAIKKMSYS GK
QTHEKWQDILKEVKFLRQLKHPNTIEYKG CYLKEHTAWLVM EYCLGSASDLLEVHKKPLQ
EVEIAAITHGALHGLAYLHSHALIHRDIKAGNILLTEPGQVKLADFGSASMASPANSFVG
TPYWMAPEVILAMDEGQYDGKVDIWSLGITCIELAERKPLFN MNAMSALYHIAQND SPT
LQSNEWTDSFRRFVDYCLQKIPQERPTSAELLRHD FVRRDRPLRVLIDL IQRTKDAVREL
DNLQYRKM KKILFQETRNGPLNESQEDEEDSEHGTSLNREMDSLGSNHSIPMSVSTGSQ
SSVNSM QEVMD ESSSELVMMHD DESTINSSSVHKKDHVFTRDEAGHGDPRPEP RPTQ
SVSQALHYRN RERFATIKSASL VTRQIHEHEQENELREQMSGYKRMRRQHQKQLIALEN
KLKAEMDEHRLKLQKEVETHANSSIELEK LAKKQVAII EKEAKVAAADEKKFQQQILA Q

Fig. 8D
29/76

QKKDLTTFLESQKKQYKICKEKIKEEMNEDHSTPKKEKQERISKHENLQHTQAEEE AHL
 LTQQRLYYDKNCRFFKRKIMIKRHEVEQQNIREELNKKRTQKEMEHAMLIRHDESTRELE
 YRQLHTLQKLRLMDLIRLQHQTELENQLEYNKRRERELHRKHVMGLRQQPKNLKAMEMQIK
 KQFQDTCKVQTQKQYKALKNHQLEVTPKNEHKTLKDEQTRKLAIALAEQYEQSINEMM
 ASQALRLDEAQEAECQALRLQLQQEMELLNAYQSKIKMQTEAQHERELQKLEQRVSLRRA
 HLEQKIEEEAALQKERSERIKNLLERQEREIETFDMESLRMGFGNLVTLDFPKEDYR

SEQ ID NO: 23 SULU3 human Nterm=missing kin partial=1-66 A=67-215
 coiled-coil1=216-425 B=426-539 coiled-coil2=540-786 Ctail=687-786
 IE LAERKPPLFNMMNAMSALYHIAQNESPTLQSNEWSDYFRNFVDSCLQKIPQDRPTSEEL
 LKHIFVLRERPETVLIDLQRTKDAVRELDNLQYRKMKLLL FQEAHNGPAVEAQEEEEQ
 DHGVGRTGTVNSVGSNQSIPSMSISASSQSSSVNSLPDVSDKSELDMMEGDHTVMSNSS
 VIHLKPEEENYREEGDPRTRASDPQSPPQVSRRHKSHYRNREHFATIRTASLVTRQMQEHE
 QDSELREQMSGYKRMRRHQKQLMTLENKLKAEMDEHRLRLDKDLETQRNNFAAEMEKLI
 KKHQAAMEKEAKVMSNEEKKFQQHIQAAQKKELNSFLESQKREYKLRKEQLKEELNENQS
 TPKKEQEWLSKQKENIQHFQAEAEANLLRRQRQYLELECRRFKRRMLLGRHNLEQDLVR
 EELNKRQTQKDLEHAMLLRQHESMQELEFRHLNTIQKMRCELIRLQHQTELTNQLEYNKR
 RERELRRKHMVEVRQQPKSLKSKEIQQKQFQDTCKIQTQYKALRNHLLETTPKSEHKA
 VLKRLKEEQTRKLAIALAEQYDH SINEMLSTQALRLDEAQEAECQVLKMQLQQELELLNAY
 QSKIKMQA EQHDRELRELEQRVSLRRALLEQKIEEMLALQNERTERIRSLLERQAREI
 EA FDSESMRLGF SNMVLSNL SPEAFSHSYPGASGW SHNPTGGPGPHWGH PMGGPPQAWGH
 PMQGGPQPWGHPGPMQGVPRGSSMGVRNSPQALRRTASGGRTEQGMSRSTS VTSQISNG
 SHMSYT

SEQ ID NO: 24 SULU3 murine Nterm=1-25 kin=26-273 A=274-422
 cc1=423-632 B=633-748 cc2=missing
 MPSTNRAGSLKDPEIAELFFKEDPEKLFTDLREIGHSGF GAVYFARDVRTNEVVAIKKMS
 YSGKQSTEKWQDIKEVKFLQRIKHPNSIEYKG CYLREHTAWLVMEYCLGSASDLLEVHK
 KPLQEVEIAAITHGALQGLAYLHSHTMIHRDIKAGNILLTEPGQVKLADFGSASMASPAN
 SFVGTPYWMAPEVILAMDEGQYDGKVDVWSLGITCIELAERKPPLFNMMNAMSALYHIAQN
 ESPTLQS NMNDSC LQKIPQDRPTSEELLKHMFLRERPETVLIDLQRTKDAVRELDNLQ
 YRKMKLLL FQEAHNGPAVEAQEEEEQDHGVGRTGTVNSVGSNQSIPSMSISASSQSSSV
 NSLPDASDDKSELDMMEGDHTVMSNSSVIHLKPEEENYQEEGDPRTRASDPQSPPQVSRRH
 KSHYRNREHFATIRTASLVTRQMQEHEQDSELREQMSGYKRMRRHQKQLMTLENKLKAE
 MDEHRLRLDKDLETQRNNFAAEMEKLIKHHQAAMEKEAKVMANEEKFQQHIQAAQKKEL
 NSFLESQKREYKLRKEQLKEELNENQSTPKKEKQEWLSKQKENIQHFQAEAEANLLRRQR
 QYLELECRRFKRRMLLGRHNLEQDLVREELNKRQTQKDLEHAMLLRQHESMQELEFRH
 LNTIQKMRCELIRLQHQTELTNQLEYNKRERELRRKHMVEVRQQPKSLKSKEIQQKQFQD

Fig. 8E
 30/76

TCKIQTRQYKALRNHLLETPKNEHKAI

SEQ ID NO: 26 GEK2 human N=1-33 kin=34-294 A=295-337 B=338-472 215
coiled-coil1=473-724 215 coiled-coil2=725-912

MAFANFRRILRLSTFEKRKSREYEHVRDLDPEVWEIVGELGDGAFGKVYAKNKETGA
LAAAKVIETKSEEELEDYIVEIEILATCDHPYIVKLLGAYYHDGKLWIMIEFCPGGAVDA
IMLELDRGLTEPQIQVVCRQMLEARNFLHSKRIIHDLKAGNVLMTLEGDIRLADFGVSA
KNLKTQKRDFIGTPYWMAPVVMCETMKDTPYDYKADIWSLGITLIEMAQIEPPHEL
NPMRVLLKIAKSDPPTLLTPSKWSVEFRDFLKIALDKNPETRPSAAQLLEHPFVSSITSN
KALRELVAEAKAEVMEEIEDGRDEGEEDAVDAASTLENHTQNSSEVSPPSLNADKPLEE
SPSTPLAPSQSQDSVNEPCSQPSGDRSLQTTSPPVAPGNENGLAVPVPLRKSRPVSMDA
RIQVAQEKKVVAEQGGDLSPAANRSQKASQSRPNSSAETLGGEKLANGSLEPPAQAAAPGP
SKRDSDCSSLCTSESMDYGTNLSTDLSLNKEMGSLSIKDPKLYKKTLKRTRKFVVDGVEV
SITTSKIISEDEKDEEMRFLRRQELRELRLLQKEEHRNQTQLSNKHELQLEQMHKRFEQ
EINAKKKFFDTELENLERQQKQQVEKMEQDHAVRRREEARRIRLEQDRDYTRFQEQLKLM
KKEVKNEVEKLPQRQRKESMKQKMEEHTQKKQLLDRDFVAKQKEDLEAMKRLTTNRRE
ICDKERECLMKKQELLRDREAALWEMEEHQLQERHQLVKQQLKDQYFLQRHELLRKHEKE
REQMQRYNQRMIEQLKVRQQQEKARLPKIQSEGKTRMAMYKKSLHINGGSAAEQREKI
KQFSQQEEKRKSERLQQQKHENQMRDMAQCESNMSELQQLQNEKCHLLVEHETQKLK
ALDESHNQNLLKE

SEQ ID NO: 29 PAK4 human Rac=1-51 A=52-224 Nterm=225-393

kin=394-658 Ctail=659-681 residues 13-23

(SAPQNFQHRVH)= Cdc42 /Rac-binding motif

MFRKKKKKRPEISAPQNFQHRVHTSFDPKEGKFVGLPPQWQNILDTRLRPKPVVDP
RVQLQPMKTVVRGSAMPVDGYISGLNDIQLSVISSNTLGRSPTSRRRAQSLGLGDE
HWATDPDMYLQSPQSERTDPHGLYLSCNGGTPAGHKQMPWPEPQSPRVLPNGLA
GPAEFQGASQRCLQLGACLQSSPPGASPPGTGNRHGMKA
HGSEEARPQSCLVGSATGR
PGGEGSPSPKTRESSLKRLFRSMFLSTAATAPPSSSKPGPPPQSKPNSSFRPPQ
KDNP
SLVAKAQS
LPSDQPVGTFSP
LTSDTSSPKSLRTAPATGQLPGRSSPAGSPRTWH
AQIS
TSNLYLPQDPTVAKG
LAGEDTG
VVTHEQFK
AALRMV
VVDQGD
PRL
LLDSYV
KIGEGSTGI
VCLAREKHSGRQ
QAVKM
MDLRK
QQR
RELLF
NEVV
IMRDY
QHF
NV
VEMY
KSYLV
GEEL
WVL
MEFLQGG
ALTD
IVSQV
RLNEE
QIA
TVCEA
VLQ
ALAYL
HAQ
GVI
HRDI
KSD
SILL
LDGRV
KLSDF
GFCA
QIS
KDV
PKR
SLV
GTP
YWM
APE
VIS
RSL
YATE
VDI
WSL
GIM
VIEM
VDGEPP
YFSD
SPV
QAM
KRL
RD
SPPP
PKL
KNS
HKV
SPV
LRF
MLV
RDP
QER
ATA
QEL
LDHP
FLL
C
TGL
PECL
VPLI
QLY
RKQT
STC

Fig. 8F
31/76

SEQ ID NO: 30 PAK5 human Rac A=missing Nterm partial=1-114
kin=115-379 Ctail=380-398

ASGAKLAAGRPFNTYPRADTDHPSRGAQGEPHDVAPNGPSAGGLAIPQSSSSSRPPTRA
RGAPSPGVLGPHASEPQLAPPACTPAAPAVPGPPGPRSPQREPQRVSHEQFRAALQLVVD
PGDPRSYLDNFNIKIGEGSTGIVCIATVRSSGKLAVKKMDLRKQQRRELLFNEVVIMRDY
QHENVVEMYNSYLVGDELWVVMEFLEGGALTIDVTHTRMNEEQIAAVCLAVLQALSVLHA
QGVIHRDIKSDSLTHDGRVKLSDFGFCAQVSKEVPRRKSLVGTPYWMAPELISRLPYG
PEVDIWSLGIMVIEMVDGEPPYFNEPLKAMKMIRDNLPPRLKNLHKVSPSLKGFLDRLL
VRDPAQRATAEELLKHPFLAKAGPPASIVPLMRQNRTR

Fig. 8G

SEQ ID NO: 1 STLK2 HUMAN

TAACAGCCCACCTCCTAGCCCCGGGCTACGCCGCCAGCCCAGTAACCCACTTTGTG
TGTCTCCCAGGCCCGATCGAAAAGCCTGGGAGGGCGGCCGAACCTACCCCCGGAGGGAG
GAGCCAGTCCGAACCAAGGCGCCACCGGCCGAGAAGCGGAGCGAGGCAGCATTGCCCTC
CATGGCCCACTGCCGGTGGCTGCAAGTGCCTGGATGCAGAATAACATAGCTGATCC
AGAAGAACTGTTCACAAATTAGAGCGCATTGGGAAAGGCTATTGGGAAGTTCAA
AGGAATTGATAACCGTACCCAGCAAGTCGTTGCTATTAAAATCATAGACCTTGAGGAAGC
CGAAGATGAAATAGAAGACATTCAAGAAGAAATACTGTCTTGAGTCATGTGACAGCTC
ATATGTAACAAACTATGGGTATTTAAAGGGGCTAAATTATGGATAATAATGGA
ATACCTGGCGGTGGTCAGCACTGGATCTTCTCGAGCTGGTCCATTGATGAGTTCCA
GATTGCTACCAGCTAAAGGAATTAAAAGGTCTGGACTATCTGATTAGAAAAGAA
AATTCAACCGAGACATAAAAGCTCCAATGTCTGCTCTAGAACAGGAGATGTTAACT
TGCTGATTTGGAGTTGCTGGTCAAGTACAGATACAGATTAAAAGAAATACCTTGT
GGGAACCTCCATTGGATGGCTCTGAAGTTATTCAACAGTCAGCTATGACTCAAAGC
TGACATTGGTATTGGAAATTACTGCTATTGAACAGCCAAGGGAGAGGCCACCTAACTC
CGATATGCATCCAATGAGAGTTCTGTTCTTATTCCCACAAACTCTCAACTCTTGT
TGGAGACTTACTAAGTCTTTAAGGAGTTATTGATGCTTGCTGAACAAAGATCCATC
ATTCGTCCTACAGAAAAGAACTTCTGAAACACAAATTCAATTGAAAAATTCAAAGAA
GACTTCTTATCTGAUTGAACAGTAGATCGTTAAGAGATGGAAGGCAGAAGGACACAG
TGATGATGAATCTGATTCCGAGGGCTCTGATTGGAATCTACAGCAGGGAAAACAATAC
TCATCCTGAATGGAGCTTACCAACCGTACGAAAGAACGCTGATCCAAGAAAGTACAGAA
TGGGGCAGAGCAAGATCTGTGCAAACCCCTGAGTTGTTGTCTATGATAATCACACCTGC
ATTTGCTGAACCTAACAGCAGGACGAGAATAACGCTAGCAGGAATCAGGCATTGAAGA
ACTCGAGAAAAGTATTGCTGTGGCTGAAGCCGCTGTCCCGCATCACAGATAAAATGGT
GAAGAAACTATTAAAAATTCAAAAGTGTTCAGCAGACGAATCCCCCTAAGAAACTTA
TTATTGGCTCTGTTCATATGGACCCAGAGAGCCCCACCAACCTACGTCAAGATTAAC
AATGCTTAACCATGAGCTCCATGTGCCCTTGGATCTTGCAACACTGAAGATTGGAA
GAAGCTATTAAACTATTTGTATGGCGTTATCATTATATTGAAAGGATTATTT
GTAAGGAATAACTTTAATACTATAGTTCACCTGTATTCTAGAAATGTTGAGACACCG
TTTGCTTTAAGTATCCTATTCTTAAGTTACGAGGATGAATACCTTCACATTGAA
TCTTAGTTGACTCTACAGTCATGAAACATACAGGTCTTCAAAGTCATTCTCAATATT
AGCTTTGTAATTCAAGCTCAAAAGCTTTTAAACACATGCAATT
CTAAAAATGACTATTGGGGGGAGGTGAAATAAGTCATACCTCTTAAACAGAAAATT
TAAGTAAAGTCTTTAAATGAAACCTGTAAAAGTATTGACTCTTCTACCAAGTTGGTATG
ATATTCCAGGCAGCTCAATGATTATCACATTGAGACCCCTGTGTTGAAGCATTACAGG
CAATGTACAGCAACAGAGGTACCTCTGGTGTAGTATTACATTCTTTAGGTAGA
AGAGGCAATTACCTTACATGGTTAGAAATTAAAGCAAGATCATTACCCAA

Fig. 9A

33/76

GGATAGGTGTTGGTAATGTTGAAGGAGTTAGTCTGGCTCATGTTTACATCTTCAC
 AAAATCCCATACTATCTGCTGGATTGGAGAGCCAAAAAATAAGCTGATTGTCATGTG
 ATAAATATCTGATCAACAGGTATGAATAACTAAATCAGCATATTGGCATGGTA
 ATAAATTGTCCTATAAACTATTATATTTGTTCTTCATAATTATCACTAATAAGCA
 TCAGTTGTTGTTAAAAGGATATTAAAGTGCAGCATTCTAGTCATATGAAAATAA
 CCATAGTACAGGATGATTCTGCCACACAAAGGTTAAATTAGATTGCACAGTTAATT
 CACTTATATTATGGTACTATTATGGGTGATGCCCTTTCTTTAAGCCCAGTACATA
 TATTATGCCCTGCTAAGTTCTGAACCTGGGCTGTATTCACTAGTTGAGAATTATTGAT
 ATTTAGTTTGATAGCTAATGTTAATTGTTGGATCTGCACAGTTGGTTTCACAA
 AAGTCATTAAAAAAATCGAGTAATTGTCAAATATTAAAAGAAAGATATTCTCCTGTA
 AGGAATACAGTTTAGTCAGGCCATTACATCCTCTTTAATTACATAACAG
 ATACTTGAGAAAGTTGTTGTTGATGCCAAGAAAATTCTTTATTGGTGCCTAT
 ATTGTAACAATTATTTAATGCATTGTATTGAAAGTAACGGTCAGTTAAATTTCAC
 CCTGCTGTGTAACGAAACACAATTACAGTTATAATCATCTGAGAAGTCTGGAGATAA
 TTTGCAACTCATGTTATGGTTAAATGAATATTGTAAGTAAAGCAACAAATT
 ATAAATTGATTATTGAAACTTACAACACAATTGCATCCAAATACAAATTGATTGCT
 TATTCAATTAGCTATTGCTCTGTAATCTGTTCTAGGTGAAGCATACTCCAGTGT
 AGGGGTTTGAAAATAAATTTAAATTTCACAGTCAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 2 STLK3 HUMAN

GACAGCAGCGCCGGCCCCGGCAGCTCCCGCGGCCCGGCCCGGCCCGGCCCGGCC
 GGCAGCACAGGCTGTCGGCTGGCCATCTGCAGGGACCGTACGAGCTGCAGGAGGTTAT
 CGGCAGTGGAGCTACTGCTGTGGTCAGGCAGCCATGCAAACCCAGGCAAGAACGTGT
 AGCAATAAAACGGATCAACTGGAAAATGCCAGACAGTATGGATGAACATTAAAGA
 AATTCAAGCCATGAGTCAGTCAGCCATCCAAACGTAGTGACCTATTACACCTCTTTGT
 GGTCAAAGATGAACCTGGCTGGTCATGAAATTACTAAGTGGAGGTTCAATGTTGGATAT
 CATAAAATACATTGTAACCGAGGAGAACACAAGAATGGAGTTCTGGAAGAGGAATAAT
 AGCAACAAATTCTAAAGAGGTTGGAGGCTTAGACTATCTACACAGAAACGGTCAGAT
 TCACAGGGATTGAAAGCTGGTAATATTCTCTGGGTGAGGATGGTCAGTACAATAGC
 AGATTGGGGTAAGTGCCTCTAGCAACAGGGGTGATGTTACCGAAATAAGTAAG
 AAAAACATTGTTGGCACCCATGTTGGATGGCTCTGAAAGTCATGGAACAGGTGAGAGG
 CTATGACTCAAGGCTGACATGTGGAGTTGGAAATAACTGCCATTGAATTAGCAACAGG
 AGCAGGCCCTATCAAATATCCTCCATGAAAGTGTAAATGTTGACTTGCACAAATGA
 TCCACCCACTTGGAAACAGGGTAGAGGATAAAGAAATGATGAAAAGTACGGCAAGTC
 CTTAGAAAATTACTTCACTGTCCTCAGAAAGATCCTCCAAAGGCCACAGCAGC
 AGAACTTTAAAATGCAAATTCTCCAGAAAGCAAGAACAGAGAGTACCTGATTGAGAA
 GCTGCTTACAAGAACACCAGACATGCCAAAGAGCCAAAAGGTAAGAAGAGTTCTGG

Fig. 9B
 34/76

GTCAAGTGGTCACCTTCAATAAACCAGAAGACGGGACTGGGAGTGGAGTGACGACGAGAT
 GGATGAGAAGAGCGAAGAAGGGAAAGCAGCTTTCTCAGGAAAAGTCACGAAGAGTAA
 AGAAGAAAATCCAGAGATTGCAGTGAGTGCAGCACCATCCCCGAACAAATACAGTCCCT
 CTCTGTGCACGACTCTCAGGGCCCACCCAAATGCTAATGAAGACTACAGAGAAGCTTCTC
 TTGTGCCGTGAACCTCGTTGAGATTAAGAAACTCCAGAAAGGAACCTTAATGACATACG
 ATTTGAGTTACTCCAGGAAGAGATACAGCAGATGGTGTATCTCAGGAGCTTCTCTGC
 TGGCTTGGTGGATGGTACCGATGTAGTTAGTGGCTGCTAATTACAGAAAGATTGTAGA
 TGATCCCAAAGCTTAAAAACATTGACATTTAAGTGGCTTCTGGCTGTGATGGTCGGA
 GATTCCCTGATGAAGTGAAGCTGATTGGGTTGCTCAGTTGAGTGTAGCTGATGTATGTC
 CCTTGATGTCACCCCTGATCTGTATGCCACCGCCACCCCTACTCCCTCAACCCCTCCC
 TCTTCTGCCATTCTCCACCCCTCACTCCATTCTAGCAAATCAGAAAGATTG
 TGAAGAGGCCGGCTTCAACAAAATGGATAAAAAAATAATTTTAAAACCTACAACACT
 CCGAGTTCTGCTTATTCTTAGCAATCCACAGTACAAGAACAGCAAATGCCACAGCTG
 CACGACTGTTGCTCATTTTCAAAAGCTATTAATATTCTAGCAATCAATTGGATAT
 CCCTTAAGTGAAAAGAATCTGAAATAACACTCAGGTGGCTTATTATTGGCAACAAAAGG
 AATTCTATCCAGAAGCCTATTCTCTTCTATTGTTATTCTGTTATAACTTT
 AATTGTACATCTGACAATACTGCCTCTTTATGTTGTATTAGAAATTAAATACTTATA
 AAATTAAAGATTATTAGCAAACCTGAATTCTAGTTAAAACGTGACTGTGAATTATT
 TTTCATATATTATGCATTACACACCTAGCTATAAGAAAAAAAGGGTTTGTATTATG
 CTTCTTGAGTTAATCTGTTATTAAACAAAAGTTGGGTCTATCTTGGAGTATT
 GTAACCTCTAAATTGAAATGACTGAATTAGGAATTGGATGCTTATTCTTGTCTG
 TTTGCCTAAAACCAATTACAATCTGACTGTCTTGGAGAGGGAGGTGCCTGCAA
 CTTTCACATTAAGAATGTGCCTGAGGCTGCTTACTCTGGAATAGTCAGATCTAA
 TTCCTCTATATAAGGTGGCATATGTTAAGTTGCTTCTATTGGACCCTAGAATGCTAT
 GTAAAATGTTGCCATTCTGTTAGATTGCTAACTATATACCCATCTGATTGGCTCTCC
 TTAAGTGTAGGATTGTTATTCTAAAGGTGATAAAACTGAAAATATCAGAATCTGAGTT
 TTACTTGAAATTTCAGAATACCCAGGTGGAGTGAAAATTGGAAGGGTTTGTCAATG
 ACTAAAAGGTAAAACGCTGTTAAGGTTCAAGAATCAATACTTCAACCCAAAGTAGCCCTC
 TGCTTGACTGTATATTATGAACTAGTAAACCTTAGGATTTGAAAATTGGAGTCTAATC
 TTTCAAGGAGGTGGCTCCAGGATGGTACCATGCTCTTCTAGCTAACCCCTAGATAT
 GGCAGCTTTAATGTACTTCAAAAGCAAATATATTACTAAGGAAAAAAAGTTATT
 ATAATTGCCTTGTCTAATTGTTAAGGTGTTAGAGCCATTGCTACAATTAAATGTA
 ATTTCATTCCATTCTATTGTTACACAACGATTACTCGAAGATGACTGCAAAGGTAAAAG
 GAAAATAAAAGTGTATTGCACAATGAAAAA

SEQ ID NO: 3 STLK4 HUMAN

CAAAAGTGGAGTCCTAGATGANTCTACCATTGCTACGATACTCCGAGAAGTACTGGAAGG
 GCTGGAATATCTGCATAAAANTGGACAGATCCACAGAGATGTGAAAGCTGGAAACATTCT

Fig. 9C
35/76

TNTTGGAGAAGATGGCTCAGTACAGATTCAAGACTTGGGGTTAGTGCTTTTAGCAAC
TGGTGGTGTATTACCCGAAATAAAGTGAAGAACACCTTGGCACCCCTGTTGGAT
GGCACCTGAAGTTATGGAACAGGTCCGTGGTTATGATTCAAAGCTGATATTGGAGTT
TGGATTACAGCAATTGAATTGGCTACAGGGCGGCTCTTATCATAAATATCCACCAAT
GAAGGTTTAATGCTGACACTGCAGAACGATCCTCCTTGGAAACTGGTGTCAAGA
TAAAGAAATGCTGAAAAAATATGGAAAATCATTAGAAAAATGATTTCATTGTGCCTCA
AAAAGATCCAGAAAAAGACCAACAGCAGCAGAACTATTAAGGCACAAATTTCCAGAA
AGCAAAGAATAAAGAATTCTCAAGAAAAACATTGCAGAGAGCACCAACCATTCTGA
AAGAGCAAAAAGGTTGGAGAGTACCAAGGTTCCAGTGGCGTCTCATAAGACAGAGGA
TGGAGGCTGGGAGTGGAGTGTGATGAATTGATGAAGAAAGTGAGGAAGGGAAAGCAGC
AATTTCACAACTCAGGTCTCCCCGAGTGAAGAAATCAATATCAAATTCTGAGCTCTTCC
AACAACTGATCCTGTGGTACTTTGCTCCAAGTCCAGAACAGATCTGCTCATCTACC
TCAGCCAGCTGGCAGATTGCTACACAGCCAACCTCAAGTCTCTCCCACCCACCGCAGA
GCCAGCAAAACAGCTCAGGCTTGTCTCAGGATCAGGTTACAAGAAACCAAGATCCC
AATCAGTCTAGTACTAAGATTCAAGGAAATTCCAAAAAGAACTAAATGATATTGATTG
ATTACTCCTGGGAGAGATAACAGCAGAGGGTGTCTCTCAGGAACCTATTCTGCTGGCCT
GGTCGACGGAAGGGATTAGTAATAGTGGCAGCTAATTGAGAAAATTGTGGAAGAAC
TCAGTCAAATCGATCTGTCACTTCAAACCTGGCATCTGGTGTCAAGGCTCAGATATTCC
TGATGATGGTAAACTGATAGGATTGCCAGCTCAGCATCAGCTAAACCAACCCCTGGA
AGAGGCAGGCTAAGGAGATTCCACACATGCGTATCTCTGTTGCTTCTATTGGCTAAACC
CACTACTGCCAAGAACCCAGCAACAAACCTCCGGCTAGGAGCTTAAAGTCTTATG
TTCTTCCTGCCATCATTCTCCTTCCACAGGGAAAGAAAAGTTGGATCACTAGTGGC
CAGCATCCCCAGAGTCCGTTAGTAAACTTACTTCATATGTCCTCTGTTCTCATTCT
GAGAAGTGGCCCATGTGCTCAAGGCCAGGAGGGAGATCTGTCAAGCTCATTCTGCTT
ACTCCAATGATGGCCAGGTGGAAAAGTAGCAGCTGTATCGGCTTCTCATCTGCTG
TTCCCCCACACCTGCCAGGATATGGACATCTGGATATCTCTTACCACTGAAGTAGAA
TTGATTGTTCAGCTGGAGCCAGAGAATTAAATTAAATTGTTTCTTGTACCTGATGT
GAATTCTAGCAACCTTGTAGGAAAAAGCACAGCCTCAGATGGAGGCAGCCTAAACTGT
GTTCTGTTGTTCATGGTGTCTAAGCGTTGCTGAAGCTGCTCAGGCACCC
TTCTTCATTGCTCTCCAGAAAGGGTGTAGCCTTAACCTCAGCTGGTCAAAACATC
TGACTGTAGCCGAACCTCAGCCATCAGATCCTCAAAGTGGAACTTGGATTGTTTAC
AGACAAACATCGAGTAATGGCTTGTAAATGTGAATTGCCAGAGGTGGTTTGAACAGG
AAAATCATAATTCATATCATTGGAGAAGTATTATTCAAATATCAAATTGAAGAAAAA
CTCAATCCTCCCATGAAAATCAGTTGCGCTGGCCTCCAAGTCGTGAGGAATGGTATGC
AAGGCTGAGATTCTACAGCAATAAGGAGACACACACTGGGCCAGAGAGGGCTGCCTC
TGCCTGCTCTCCTGCACTGACCCCTTGGAGGGGTCTCTGTGTGCTGAAGCTAAG
ATGGAAAGTGAACCAACATGTGCCGTGACCTTGTAGGTTATGAGTAGACAGTGT
TGATTCTACAGAAATAATAAATTCTTAGGTTAAAAAGAGCACTCATAATG

Fig. 9D
36/76

CAATATGTGAATAATCAGTGAGGTTGATTTTCTTTCTACCGTTCATAGCTTG
 TCTAACTGCTAGTAACCCTACCGAGTTTATATATGAGTGGGATACTCAATCTGGCCTTA
 AAAAGATAACAAAGATGGCTGTGGTCCCTGGAAAGGGGGAGAGTTGCCCTTACAGA
 ATCACTCGAGCCCTTCCAGCACTGTTGGTCTGATGAACAAGGTTGTTTACCTTATTT
 CTCTTGGAACATATCTGAAAACCTCCCCACAAATAACTGTCACACCTTTGTTCTT
 CTGAGTCTTAGTTAGTCATGGCTTCTCACCTGCTCTAGGTGCAAAGGCATGTTG
 GGAAAGAGATGGATGTTGGGAGGAAGAGAGGGAGATGGATTCAAGTGGAGTTAGGAGG
 AGAGTAGGTGAGATGATCAGACACCGGAGTTAACGTCCCAGCAGTCTGGTAAAAGGAG
 GGAGCCTGCTGAGCCAGGAGGGAGAAAGAAGATTGACCAGCTGCTAGAAAAATACTTA
 GCTTTCTTTCTTTGTGGAGGGGGACGGAGAGGAACAAGGATGGGAGGTAGG
 AATGAGGTATAGAAAAGAGATAGCATCTTCTTGGCACAAAGACTAGTGGCTTACCGCTTA
 CCTTAGAGTTTGTTTTTTCAAACCCATCAAAATCTACTTATTATGAATCAA
 GGGGTGGCAGCATCACTCTGTTCTAGCATTCTTGTGGAGATGGCTGGTGCCTAGCTGG
 GAGTGAGCAGCAGCCCCTGTTCACTTCTAGCCCATCATTACCTGTGAACCTGC
 AGTGGGGCAGTCATGGCAAATAGAATTGGCTGGGTTCTCCTTCTTCAGTTCTTCTG
 TTGCCCCGCTAGGAATTAGAACAGACACCATGTCCAGGACAGTGTACTTCTTCTG
 CATGATGTGTGGTAGACTCCCTTGCTGGCTTGTGAGTGAACATGAGAAAATACATGAA
 CAGAAACTGCCAGGTGGAACAGCACGTAACCTAGTGAGTGACTGTACTCCTTCTAGGA
 ATGCTGATTCAAGAGTCACCTCTTGACTAGGTCCAGGATCCCCCTGTCCCTGGAGTAG
 GGACTAACTATAGCACAAAGTAATATGTGCCATGCTATTGTGAAATGTTGGTCTTC
 TAAACGACTAAAGGATTGTTGGGTTTGCTTAAGTTGAACCAAATCTAGAGCCAG
 CTGATAATATTAATAATCTGGAGGAGAGAATAATGATGTACCAATAAGTGAGATTCT
 CCTTATGATGTATGCTAGGTTATGGAAGATGAAATATTCAACTTTCTCCTTTT
 TGGACTTGTATTTACTGCATGTTCTCATTAAATCAATAAGAGTAAATTGTCA
 AAAAAAAAAAAAAAAA

SEQ ID NO: 4 STLK5 HUMAN

CTCATCTGTACACACTCATGGATGGCATGAATGAGCTGGCGATTGCTTACATCCTGCAG
 GGGGTGCTGAAGGCCCTGACTACATCCACACATGGGATATGTACACAGGAGTGTCAA
 GCCAGCCACATCCTGATCTCTGTTGGATGGGAAGGTCTACCTGTCGGTTGCGCAGCAAC
 CTCAGCATGATAAGCCATGGCAGCGGCAGCGAGTGGTCCACGATTTCCAAGTACAGT
 GTCAAGGTTCTGCCGTGGCTAGCCCCGAGGTCTCCAGCAGAAATCTCCAGGGTTATGAT
 GCGCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAACTGGCCAACGGCATGTC
 CCCTTAAGGATATGCCCTGCCACCCAGATGCTGCTAGAGAAACTGAACGGCACAGTGCC
 TGCCCTGTTGGATACCAGCACCCTCCCGCTGAGGGAGCTGACCATGAGCCCTCGCGCTCA
 GTGGCCAACCTCTGGCTGAGTGACAGCCTGACCACAGCACCCCCGGCCCTCAAACGGT
 GACTGCCCTCCACCCCTACCAACCGAACCTCTCCCCCACTCCACCAACTTGTGGAG
 CAGTGCCTCAGCGAACCCGGATGCCAGGCCAGTGCAGCACCCCTCTGAACCAACTCT

Fig. 9E

37/76

TTCTTCAAGCAGATCAAGCGACGTGCCTCAGAGGCTTGCCCCAATTGCTCGTCCTGTC
 ACCCCCACATCACCAATTGGAGGGCAGCCAGTCTCAGGACCACAGTGGAAATTTGGCCTG
 GTAACAAACCTGGAAAGAGCTGGAGGTGGACGATTGGGAGTTCTGAGCCTCTGCAAACGT
 GCGCATTCTCCAGGCCAGGGATGCAGAGGCCACCCAGAGGCCCTTGAGGGCCGGCCAC
 ATTCCCGCCCTCCTGGCAGATTGGTAGAAAGGACATTCTCCAGGAAAGTTGACTGCT
 GACTGATTGGAAAGAAAATCCTGGAGAGATACTTCACTGCTCCAAGGCTTTGAGACAC
 AAGGGAATCTCAACAACCAGGGATCAGGAGGGTCAAAGCCGACATTCCCAGTCTGTGA
 GCTCAGGTGACCTCCTCCGAGAAGAGAGATGCTGCTCTGGCCCTGGGAGCTGAATTCCA
 AGCCCAGGGTTGGCTCTAAACCCGAGGACCACCTCTCCAGTGCTTGCACCA
 GCCTCATTCTATTAACTTGCTCTCAGATGCCTCAGATGCTATAAGTCAGTGAAAGGGC
 AAGTAGTAAGCTGCCTGCCTCCCTCCCTCAGACCTCTCCCTCATATTCCAGAGAAGGG
 CATTCTGTCTTTAAGCACAGACTAAGGCTGAAACAGTCCATCCTATTCCCTCTG
 GCTTGGGCCCTGACACCTAACAGTCTTCCACGGTTATGTGTGTGCCTCATTCCCTCC
 ACCAAGAATCCATCTAGCGCCTCCTGCCAGCTGCCCTGGTGCCTTCTCCAAGGGCCATC
 AGTGTCTTGCCTAGCTTGGGCTTAAGTCCTATGCTGTGTTAGTTGTTAGAAGC
 AAATTAAAATTTCAGAGACGCTG

SEQ ID NO: 9 ZC1 HUMAN

GAGACCATGGCGAACGACTCTCCCGGAAAGTCTGGTGGACATCGACCTCTCCCTG
 CGGGATCCTGCTGGATTGGAGCTGGTGGAAAGTGGTGGAAATGGCACCTATGGACAA
 GTCTATAAGGGTCGACATGTTAAACGGGTAGTGGCAGCCATCAAAGTTATGGATGTC
 ACTGAGGATGAAGAGGAAGAAATCAAACGGAGATAAAATGCTAAAGAAATACTCTCAT
 CACAGAAACATTGCAACATATTGGTCTTCACTAAAAAGAGCCCTCCAGGACATGAT
 GACCAACTCTGGCTTGTATGGAGTTCTGTGGGGCTGGTCCATTACAGACCTTGTGAAG
 AACACCAAAGGGAACACACTCAAAGAAGACTGGATCGCTTACATCTCCAGAGAAATCTG
 AGGGGACTGGCACATCTTACATTGATCATGTGATTCACTGGGATATCAAGGGCCAGAAT
 GTGTTGCTGACTGAGAATGCGAGGGTAAACTTGTGACTTGGTGTGAGTGCTCAGCTG
 GACAGGACTGTGGGGCGGAGAAATCGTTCACTGGGACTCCCTACTGGATGGCTCTGAG
 GTCATGCCCTGTGATGAGAACCCAGATGCCACCTATGATTACAGAAGTGATCTTGGTCT
 TGTGGCATTACAGCATTGAGATGGCAGAAGGTGCTCCCCCTCTGTGACATGCATCCA
 ATGAGAGCACTGTTCTCATTCCAGAAACCCCTCTCCCCGGCTGAAGTCAAAAAAATGG
 TCGAAGAAGTTTTAGTTATAGAAGGGTGCCTGGTGAAGAATTACATGCAGCGGCC
 TCTACAGAGCAGTTGAAACATCTTTATAAGGGATGCCAAATGAAAGGCAAGTT
 AGAATCCAGCTTAAGGATCATATAGATCGTACCAAGGAAGAAGAGAGGGCAGAAAGATGAA
 ACTGAGTATGAGTACAGTGGAGTGAGGAAGAAGAGGGAGGAGTGCCTGAACAGGAAGGA
 GAGCCAAGTTCCATTGTGAACGTGCCTGGTGAAGTCTACTCTCGCCGAGATTCTGAGA
 CTGCAGCAGGAGAACAGGAACGTTCCAGGCTTCTGGAGACAAACAGTTACTACAGGAG
 CAACAGCTCCGGAGCAGGAAGAATATAAGGCAACTGCTGGCAGAGAGACAGAAGCGG

Fig. 9F
 38/76

ATTGAGCAGCAGAAAGAACAGAGGCACGGCTAGAAGAGCAACAAAGGAGAGAGCGGGAA
GCTAGAAGGCAGCAGGAACGTGACAGCGAAGGAGAGAACAGAAGAAAAGAGGGTCTA
GAGGAGTTGGAGAGAACGGCGAAAGAACAGAGGAGAGGAGACGGCAGAACAGAAGAAAAG
AGGAGAGTTGAAAGAACACAGGAGTATATCAGGCAGCTAGAACAGAGGAGCAGCGGAC
TTGGAAGTCCTCAGCAGCAGCTGCTCCAGGAGCAGGCCATGTTACTGGAGTGCCGATGG
CGGGAGATGGAGGAGCACCGGCAGGCAGAGAGGCTCCAGAGGCAGTTGCAACAAGAACAA
GCATATCTCCTGTCTACAGCATGACCAGTAGGAGGCCACCCGAGCACTCGCAGCAG
CCGCCACCACCGCAGCAGGAAAGGAGCAAGCCAAGCTCCATGCTCCGAGGCCAAAGCC
CACTACGAGCCTGCTGACCGAGCGAGAGGTGGAAGATAAGATTAGGAAAACTAACCAC
AGCTCCCCCTGAAGCCCAGTCTAACGAGACAGGCAGAGTATTGGAGCCACCAAGTGCCTTCC
CGATCAGAGTCTTTCCAATGGCAACTCCGAGTCTGTGCATCCGCCCTGCAGAGACCA
GGGAGCCACAGGTTCTGTGAGAACACATCTCGCTCCCTGTTCTGTCCCGTCAGAT
TCCCCACTGCAGGGCAGTGGGAGCAGCAGAACAGCCAGGACAGAGAAACCTCACCAGT
ATTGAGCCAGGCTTCTGTGGAGAGAGTGGAGAACGCTGGTGCCAGACCTGGCAGTGGC
AGCTCCTCAGGGTCCAGCAACTCAGGATCCAGCCCAGCCGGTCTCACCCGGTCTCAGAGT
GGCTCCGGGAACGCTTCAGAGTGGAGATCATCATCCAAGTCTGAAGGCTCTCCATCTCAG
CGCCTGGAAAATGCAGTGAACACCTGAAGATAAAAGGAAGTTTCAGACCCCTCAAG
CCTGCTGATCTGACCGCACTGGCAAAGAGCTTCAGCAGTGGAGATGTACGGCACCT
CACAAAGTAACGGACTACTCCTCATCCAGTGGAGCTGGGAGCAGGATGAGGAGGAC
GACGATGTGGAGCAGGAAGGGCTGACGAGTCCACCTCAGGACCAGAGGACACCAGAGCA
GCGTCATCTCTGAATTGAGCAATGGTAAACGGAATCTGTGAAAACCATGATTGTCCAT
GATGATGTAGAAAGTGGCCGGCATGACCCCATCCAAGGAGGGACTCTAACGCTCCGC
CGGACTCAGTCCCTAGTAGCACACTCCAGAAACACAAATCTCCTCCTTACACCT
TTTATAGACCCAGATTACTACAGATTCTCCATCTAGCGGAACACAGTGAATCTGTG
GTGGGATTTCTGTGATGGAGAGACCAGAACGCCATAAGGCAAGATCTACCCGGAAA
GGCTCAGGGTCAATGTGAATCCTACCAACACTAGGCCACAGAGTGACACCCGGAGATT
CGTAAATACAAGAAGAGGTTAACCTGAGATTCTGTGCTGCCATTGGGAGTGAAAT
TTGCTAGGGTACAGAGAGTGGCTGATGCTGGACAGAACAGTGGCAAGGAAAGGTC
TATCCTCTTATCAACCGAACGACGATTCAACAAATGGACGTACTTGAGGGCTTGAATGTC
TTGGTACAATATCTGGAAAAAGGATAAGTTACGTGTTACTATTGTCCTGGTTAAGA
AATAAAATACCTACAATGATCCAGAACAGTTGAGAACAGCAGGGATGGACAACCGTAGGG
GATTGGAGGATGTGACATTATAAGTTGAAAGAACATCAAATTCTGGT
ATTGCTTGAAGAGTTCTGTGAAGTCTATCGTGGCACCAAGCCATATCACAAATT
ATGGCCTTAAGTCATTGGAGAATTGGTACATAAGCCATTACTGGTGGATCTCACTGTT
GAGGAAGGCCAGAGGTTGAAAGTGATCTATGGATCCTGTGCTGGATTCCATGCTGTTGAT
GTGGATTCAAGGATCAGTCTATGACATTATCTACCAACACATATCCAGTGTAGCATCAA
CCCCATGCAATCATCATCCTCCCAATACAGATGGAATGGAGCTCTGGTGTGCTATGAA
GATGAGGGGTTATGTAACACATATGGAAGGATACCAAGGATGTAGTTCTACAGTGG

Fig. 9G
39/76

GGAGAGATGCCTACATCAGTAGCATATATTGATCCAATCAGACAATGGGCTGGGGAGAG
 AAGGCCATAGAGATCCGATCTGTGAAACTGGTCACTTGGATGGTGTTCATGCACAAA
 AGGGCTCAAAGACTAAAATTCTTGTGTGAACGCAATGACAAGGTGTTCTTGCCTCTGTT
 CGGTCTGGTGGCAGCAGTCAGGTTATTCATGACCTTAGGCAGGACTTCTTCTGAGC
 TGGTAGAAGCAGTGTGATCCAGGGATTACTGGCCTCCAGAGTCTTCAAGATCCTGAGAAC
 TTGGAATTCCTTGTAACT

SEQ ID NO: 10 ZC2 HUMAN

GCTTCGGGGAGGTCTATGAGGGTCGTATGTCAAAACGGGCCAGCTGCAGCCATCAAG
 GTTATGGATGTACAGGGATGAAGAGGAAGAAATCAAACAAGAAATTACATGTTGAAG
 AAATATTCTCATCACCGGAATATTGCTACATACTATGGTGTCTTATCAAAAAGAACCCA
 CCAGGCATGGATGACCAACTTGGTGTGATGGAGTTTGCTGGCTCTGTCACC
 GACCTGATCAAGAACACAAAAGGTAACACGTTGAAAGAGGAGTGGATTGCATACATCTGC
 AGGGAAATCTTACGGGGCTGAGTCACCTGCACCAGCATAAAGTGATTATCGAGATATT
 AAAGGGAAAATGTTGCTGACTGAAAATGCAGAAGTTAAACTAGTGGACTTGGAGTC
 AGTGCTAGCTTGTATCGAACAGTGGCAGGAGGAATACTTCATTGAACTCCCTACTGG
 ATGGCACCAAGAAGTTATTGCCTGTGATGAAAACCCAGATGCCACATATGATTCAAGAGT
 GACTTGTGGTCTTGGTATCACCAGCATTGAAATGGCAGAAGGTGCTCCCCCTCTGT
 GACATGCACCCATGAGAGCTCTTCTCATCCCCGGATCCAGCGCTCGGCTGAAG
 TCTAAGAAGTGGTAAAAAAATTCCAGTCATTATTGAGAGCTGCTTGGTAAAGAACATAC
 AGCCAGCGACCAGCAACAGAACATTGATGAAGCATCCATTATACGAGACCAACCTAAT
 GAGCGACAGGTCCGATTCAACTCAAGGACCATATTGATAGAACAAAGAACAGCAGGAG
 GAAAAAGATGAGACAGAGTATGAGTACAGTGGAAAGTGAGGAAGAACAGGAGGAGAACATGAC
 TCAGGAGAGCCCAGCTCCATCCTGAATCTGCCAAGGGAGTCACGCTGCGGAGGGACTT
 CTGAGGCTGCAGCTGCCAACAGGAGCGTTCTGAGGCCACTGGAGGCAGCAGCTGGAG
 CAGCAGCAGCGGGAGAATGAGGAGCACAAGCGGAGCTGCTGGCCAGCGTCAGAGCGC
 ATCGAGGAGCAGAACAGAGCAGAGGGAGCTGGAGGAGCAACAAAGCGAGAGAACAG
 CTGCGGAAGCAGCAGGAGAGGGAGCAGCGCCGGCACTATGAGGAGCAGATGCGCCGGAG
 GAGGAGAGGAGGCGTGCAGCATGAACAGGAATATAAGCGCAAACAAATTGGAAGAACAG
 AGACAAGCAGAAAGACTGCAGAGGCAGCTAAAGCAAGAAAGAGACTACTTAGTTCCCTT
 CAGCATCAGCGGCAGGAGCAGAGGCCCTGTGGAGAACAGGCACTGTACCAATTAAAGAA
 GGAATGAGTCTTAGTGAGAACGCCAGCATGGCCAAGGAGGAGAACAGGTCAAGGCTC
 AACCGGCAAAGTCCCTGCCATGCCTCACAGGTTGCCAACAGGATATCTGACCCAAC
 CTGCCCTAACGGTCGGAGTCCTCAGCATTAGTGGAGTTAGCCTGCTGAACACCCCC
 ATGCTCAGACCAGTCGATCCCCAGATCCCACATCTGGTAGCTGAAAATCCCAGGGACCT
 GCCTTGACCGCCTCCAGTCAGTCACGAGCAGCCCACAAAGGGCTCTGGGTTTCAAG
 GAGGCTCTGAACGTGACCTCCACCGCGTGGAGATGCCACGCCAGAACTCAGATCCCACC
 TCGGAAAATCCTCCTCTCCCCACTCGCATTGAAAAGTTGACCGAACGCTTTGGTACGA

Fig. 9H
 40/76

CAGGAAGAAGACATTCCACCAAAGGTGCCTCAAAGAACAACTTCTATATCCCCAGCATT
 GCCAGAAAGAATTCTCCTGGGAATGGTAGTGCTCTGGGACCCAGACTAGGATCTCAACCC
 ATCAGAGCAAGCAACCCCTGATCTCCGGAGAACTGAGCCCCTCTTGGAGAGCCCCCTTCAG
 AGGACCAGCAGTGGCAGTCCCTCAGCTCCAGCACCCCTAGCTCCCAGGCCAGCTCCCAA
 GGAGGCTCCAGCCTGGATACAAGCAGGATCCAGTGAACGCACCAGAGTTGAGCCAAC
 AGTAAGTCAGAAGGGATCACCTGTGCTCCCCATGAGCCTGCCAAGGTGAAACCAGAAGAA
 TCCAGGGACATTACCCGGCCAGTCGACCAGCTAGCTACAAAAAAAGCTATAGATGAGGAT
 CTGACGGCATTAGCCAAAGAACTAAGAGAACTCCGGATTGAAGAAACAAACCGCCCAATG
 AAGAAGGTGACTGATTACTCCTCCAGTGAGGAGTCAGAAAGTAGCGAGGAAGAGGAG
 GAAGATGGAGAGAGCGAGACCCATGATGGGACAGTGGCTGTCAGCGACATAACCAAGACTG
 ATACCAACAGGAGCTCCAGGCAGCAACGAGCAGTACAATGTGGGATGGTGGGACGCAT
 GGGCTGGAGACCTCTCATGCGGACAGTTCAAGCAGTATTCAAGAGAAAGGAACCTTG
 ATGATTAGAGAGACGTCTGGAGAGAAGAAGCGATCTGGCCACAGTGAACAGCAATGGCTT
 GCTGGCCACATCAACCTCCCTGACCTGGTCAGCAGGCCATTCTCAGCTGGAACCCCG
 ACTGAGGGACTGGGGCGCTCAACCCATTCCAGGAGATGGACTCTGGACTGAATAT
 GGCATGGGGAGCAGCACAAAGCCTCTCACCCCTTGAGGACCCAGAGTATACCAAG
 ACGTCTCCACTGATGAAGATGAAGAGGATGAGGAATCATCAGCCGAGCTCTGTTACT
 GGCAGACTCTTAGGCAAGAACAGGCCAACTCAATGAAGCAAGAAAGATTCTGGTGT
 AATGTAAACCCAAACACATTGGCCTCATAGCAGACACACCAGAAATCAGAAAATACAAG
 AACCGATTCAACTCAGAAATACTTTGTGAGCTCTGTGGGTGAAACCTCTGGTGGG
 ACTGAAAATGGCCTGATGCTTTGGACCGAAGTGGCAAGGCAAAGTCTATAATCTGATC
 AACCGGAGGCATTTAGCAGATGGATGTGCTAGAGGACTGAATGTCCTGTGACAATT
 TCAGGAAAGAAGAATAAGCTACGAGTTACTATCTTCACTGGTTAAGAAACAGAATACTA
 CATAATGACCCAGAAGTAGAAAAGAACAAAGGCTGGATCACTGTTGGGACTTGGAAAGG
 TGATGCTACATTATAAGTTAAATATGAAAGGATCAAATTGGTGTGATTGCCTTAAAG
 AATGCTGTGAAATATGCTTGGCTCTAAACCGTATCATAAATTGTCATGGCATTAAAG
 TCTTTGAGATCTCCAGCACAAGCCTCTGCTAGTTGATCTCACGGTAGAAGAAGGTCAA
 AGATTAAAGGTATTTGGTACACACTGGTTCCATGTAATTGATGTTGATTAGGA
 AACTCTTATGATATCTACACACCCTCATATTAGGGCAATATCACTCCTCATGCTATT
 GTCATCTTGCTAAACAGATGGAATGAAATGCTTGTGCTATGAGGATGAGGGGTG
 TATGTAACACCTATGGCGGATAACTAAGGATGTGGTGTCCAATGGGAGAAATGCC
 ACGTCTGTGGCCTACATTCCAATCAGATAATGGCTGGGCGAGAAAGCTATTGAG
 ATCCGGTCAGTGGAAACAGGACATTGGATGGAGTATTTATGCTAAGCGAGCTAAAGG
 TTAAAGTTCTATGTGAAAGAAATGATAAGGTATTTTGATCCGTGCGATCTGGAGGA
 AGTAGCCAAGTGTGTTTCTAGACCCCTAACAGAAATTCCATGATGAACTGGTAACAGAAG
 AGCACTGGCACTTATCTCATGGCGTTATTCATAATTAAAAGAACATAACTCATGTGG
 ACTTATGCCAGTCTAGAGGCAGAATCAGAAGGCTGGTTGAACATATCGCTTCCCTTT
 TCCTCTCCCTCCGCCCCCTCCAGTACAGTCCATCT

Fig. 9 I
 41/76

SEQ ID NO: 11 ZC3 HUMAN

GCATTTGGGGAGGTGTATGAGGGTGGCATGTCAAGACGGGGCAGCTGGCTGCCATCAAG
 GTCATGGATGTCACGGAGGACGAGGAAGAGATCAAACAGGAGATCAACATGCTGAAA
 AAGTACTCTCACCAACCGAACATGCCACCTACTACGGAGCCTCATCAAGAAGAGCCCC
 CCGGGAAACGATGACCAGCTCTGGCTGGTGTGGAGTTCTGTGGTGTGGTCAGTGACT
 GACCTGGTAAAGAACACAAAAGGCAACGCCCTGAAGGAGGACTGTATGCCCTATATCTGC
 AGGGAGATCCTCAGGGTCTGGCCATCTCCATGCCACAAGGTGATCCATCGAGACATC
 AAGGGGAGAACATGTGCTGTCAGAGAACATGCTGAGGTCAAGCTAGTGGATTGGGGTG
 AGTGCTCAGCTGGACCGCACCGTGGCAGACGGAACACTTCATTGGACTCCCTACTGG
 ATGGCTCCAGAGGTATGCCCTGTGATGAGAACCTGATGCCACCTATGATTACAGGAGT
 GATATTGGTCTCTAGGAATCACAGCCATCGAGATGGCAGAGGGAGCCCCCTGTGT
 GACATGCACCCATGCGAGCCCTTCTCATTCCCTGGAACCCCTCCGCCAGGCTCAAG
 TCCAAGAAGTGGCTAAGAACATTGACTTCATTGACACATGTCTCATCAAGACTTAC
 CTGAGECGCCCACCCACGGAGCAGCTACTGAAGTTCATTCCCTCATCCGGGACCAGCCCACG
 GAGCGGCAGGTCCGCATCCAGCTTAAGGACACATTGACCGATCCGGAAAGAAGGGGGT
 GAGAAAGAGGAGACAGAACATGAGTACAGCGCAGCGAGGAGGAAGATGACAGCCATGGA
 GAGGAAGGGAGAGCCAAGCTCCATCATGAACGTGCCTGGAGAGTCGACTCTACGCCGGAG
 TTTCTCCGGCTCCAGCAGGAAATAAGAGCAACTCAGAGGCTTAAAACAGCAGCAGCAG
 CTGCAGCAGCAGCAGCAGCACACATCAAACACCTGCTGCACCAGCGG
 CAGCGGCGCATAGAGGAGCAGAAGGAGGAGCGGGCGCCGCGTGGAGGGAGAACAGCGGCGG
 GAGCGGGAGCAGCGGAAGCTGCAGGAGAACGGAGCAGCAGCGGCGGCTGGAGGGACATGCAG
 GCTCTCGGGGGAGGAGCAGCGGAGCAGCGAGCAGGAATATATTGTCAC
 AGGCTAGAGGAGGAGCAGCAGCTGAGATCCTCAGCAACAGCTGCTCCAGGAACAG
 GCCCTGCTGCTGGAATAAGCGGAAGCAGCTGGAGGAGCAGCGGAGTCAGAACGTCTC
 CAGAGGAGCTGCAGCAGGAGCATGCCTACCTCAAGTCCCTGCAGCAGCAGAACAGCAG
 CAGCAGCTCAGAAACACAGCAGCAGCTCCTGCCTGGGGACAGGAAGCCCTGTAC
 CATTATGGTGGGGCATGAATCCGCTGACAAACCAGCCTGGGCCAGAGGGTAGAAGAG
 AGAACAAAGGATGAACAAAGCAGCAGAACTCTCCCTGGCCAAGAGCAAGCCAGGCAGCAG
 GGGCCTGAGCCCCCATCCCCAGGCCTCCCCAGGGCCCCAGGACCCCTTCCAGACT
 CCTCCTATGCAGAGGCCGGTGGAGCCCCAGGAGGGACCGCACAAAGAGCCTGGTGGCACAC
 CGGGTCCCACTGAAGCCATATGCAGCACCTGTACCCGATCCAGTCCCTGCAGGACCAAG
 CCCACCCGAAACCTGGCTGCCTCCAGCCTCCATGACCCGACCCCTGCCATCCCCGA
 CCACACTGCCACGCCAGTCCCAGGGAGCTGTACCGCCAGAATTCAAGACCCACCTCT
 GAAGGACCTGGCCCCAGCCGAATCCCCAGCCTGGTCCGCCAGATAACGAGGCCCA
 CCCAAGGTGCCTCAGAGGACCTCATCTATGCCACTGCCCTTAACACCAAGTGGGGCCGGA
 GGGTCCCGGCCAGGCCAGGCAGTCCGTGCCAGACCTCGCAGCAACTCCGCTGGCAAATC
 TATCTGCAAAGGCAGGGCAGAGCAGGGCACCCAAAGCCTCAGGGCCCCCTGCTCAGCC
 CCTGGCCCGCCAAACGCCCTAGTAACCCGACCTCAGGAGGAGCAGCCCTGGCTGGAA

Fig. 9J
42/76

CGCTCGGACAGCGTCTTCCAGCCTCTCACGGCACCTCCCCAGGCTGGCTCACTGGAG
 CGGAACC CGGTGGGAGTCTCCTCAAACCGAACAGCTCCCTGTGCTCTCCCTGGGAAT
 AAAGCCAAGGCCGACGACCACCGCTCACGGCCAGGCCGGCCGCAGACTTTGTGTTGCTG
 AAAGAGCGGACTCTGGACGAGGCCCTGGCCTCCAAGAAGGCCATGGACTACTCGTCG
 TCCAGCGAGGGAGGTGGAAAGCAGTGAGGACGACGAGGAAGGCGAAGGCCGGCCAGCA
 GAGGGGAGCAGAGATAACCCCTGGGGGCCGCATGGGATACAGACAGCGTACGCCATG
 GTGGTCCACGACGTCGAGGAGATACCGGGACCCAGCCCCATACGGGGCGGCACCATG
 GTGGTCCAGCGCACCCCTGAAGAGGAGCGGAACCTGCTGCATGCTGACAGCAATGGGTAC
 ACAAACCTGCCTGACGTGGTCCAGCCCAGCCACTCACCCACCGAGAACAGCAAAGGCCAA
 AGCCCACCCCTGAAGGATGGGAGTGGTACTACCAAGTCTCGTGGGCTGGTAAAGGCCCT
 GGCAAGAGCTCGTTACGATGTTGTGGATCTAGGGATCTACCAGCCTGGAGGCAGTGGG
 GAAGCATCCCCATCACAGCCCTAGTGGGTGGAGAGGGCACTCGGCTCGACCGAGCTGCA
 TACGACGTGAGGAAGGGTCTGTGGTCAACGTGAATCCCACCAACACCCGGGCCACAGT
 GAGACCCCTGAGATCCGGAAGTACAAGAACGATTCAACTCCGAGATCCTCTGTGCAAGC
 CTTGGGGGGTCAACCTGCTGGTGGCACGGAGAACGGCTGATGTTGCTGGACCGAAGT
 GGGCAGGGCAAGGTGTATGGACTCATTGGCGGCACGCTTCAGCAGATGGATGTGCTG
 GAGGGGCTCAACCTGCTCATCACCATCTCAGGGAAAAGGAACAAACTGCGGGTGTATTAC
 TTGTCCTGGCTCCGGAACAGATTCTGCACAATGACCCAGAAGTGGAGAACAGCAGGGC
 TGGACCACCGTGGGGGACATGGAGGGCTGCGGGCACTACCGTGTGAAATACGAGCGG
 ATTAAGTTCTGGTCATGCCCTCAAGAGCTCCGTGGAGGTGTATGCCCTGGCCCCAAA
 CCCTACCACAAATTATGGCCTTCAAGTCCTTGCCGACCTCCCCACCGCCCTCTGCTG
 GTCGACCTGACAGTAGAGGGAGGGCAGCGGCTCAAGGTATCTATGGCTCCAGTGTG
 TTCCATGCTGTGGATGTCGACTCGGGAACAGCTATGACATCTACATCCCTGTGACATC
 CAGAGCCAGATCACGCCCATGCCATCATCTTCCTCCCCAACACCGACGGCATGGAGATG
 CTGCTGTGCTACGAGGACGGGGTCTACGTCAACACGTACGGCGCATATTAGGAT
 GTGGTGTGCTGAGTGGGGGAGATGCCTACTTCTGTGGCTACATCTGCTCCAACCAGATA
 ATGGGCTGGGTGAGAACGCCATTGAGATCCGCTCTGTGGAGACGGGCCACCTGACGGG
 GTCTTCATGCACAAACGAGCTCAGAGGCTCAAGTTCTGTGAGCGGAATGACAAGGTG
 TTTTTGCCTCAGTCGCTCTGGGGCAGCAGCCAAGTTACTTCATGACTCTGAACCGT
 AACCGCATCATGAACCTGGTACGGGCCCTGGCTGGGCTGTCCCACACTGGACCCAGC
 TCTCCCCCTGCAGCCAGGCTCCGGGCCCTCTTCCCCCTCCCTGGCTTTGCTT
 TTACTGGTTGATTCACTGGAGCCTGCTGGAACGTGACCTCTGACCCCTGA

SEQ ID NO: 12 ZC4 HUMAN

CAATGTTAACCCACTCTATGTCTCTGCATGTAAAAAACCACTAACATCCACATGTATGA
 AAAGGAGTTCACTCTGAGATCTGCTGTGGTTCTTGTGGGAGTCAATTGCTGTTGGG
 AACCCGATCTAATCTATATCTGATGGACAGAAGTGGAAAGGCTGACATTACTAAACTTAT
 AAGGCAGAACGACATTCCGCCAGATTCAAGTCTTAGGCCACTCAATTGCTGATTACCAT

Fig. 9K
43/76

CTCAGGTATAAGAACAGACTTCGGGTATCATCTGACCTGGTTGAGGAACAAGATTT
 GAATAATGATCCAGAAAGTAAAAGAAGGCAAGAAGAAATGCTGAAGACAGAGGAAGCCTG
 CAAAGCTATTGATAAGTTAACAGGCTGTGAACACTTCAGTGTCTCCAACATGAAGAAC
 AACATATATTGCAATTGCTTGAAATCATCAATTCACCTTATGCATGGCACCAAAGTC
 CTTGATGAAAGCACTGCTATTAAAGTATTCCAACACTTGATCATAAGCCAGTGACAGT
 TGACCTGGCTATTGGTTCTGAAAAAAGACTAAAGATTTCTCAGCTCAGCAGATGGATA
 TCACCTCATCGATGCAGAATCTGAGGTTATGTCTGATGTGACCTGCCAAGAACATCCCCT
 GGAAATCATTATAACCACAGAACATCATCATTACCTGATTGCTTGGAAATTGGCATGAT
 GCTCACCTCAATGCTGAAGCCCTCTGTGGAGCAAATGAACAACACTTTCAAGAACAT
 CCTTGAATGTGGAAAGACATACCATCTTCTATAGCTTTGAATGTACACAGCGAACAC
 AGGATGGGGCCAAAAGGCCATTGAAGTGCCTCTTGCAATCCAGGGTTCTGGAAAGTGA
 GCTGAAGCGCAGGTCAATTAGAACAGCTGAGGATTCTGTGCACCCGGGGTACAAGCTGTT
 CTTACCTCTACCCCTGCAGAATCACCACAGCCGGTTACTTCATGACACTGGAAAAT
 TGAAGAGCTCAAAGCAATTATGATGTCTAAAGTTCCAGTGATTATTACACATTAT
 AAACATCATGTATAGGCAGTCTGCATCTTCAGATTCAAGAGATTAAATGAGTATTGAGT
 TTATTTTAGTAAAGATTAAATCCAAAATTTACTTTAATGTAGCACAGAACATAGTTTA
 ATGAGAAATGCAGCTTATGTATAAAATTAACATAGCAAGCTCTAGGTACTCCAATGGT
 GTACAATGTCTTGACAAAACTTGTAACCTTGTTACTGTGAATTCAAACATTACTCT
 TTGGACAGTTGGACAGTATCTGTATTCAAGATTACATGGAGTAAAGAACCTGTT
 ATGAATTAGATTACAAGCAGCCTCAAAAGAATTGGCACTGGATAAGATTTTCAAGAAA
 AGAAAAACATCGGCAAAC

SEQ ID NO: 17 KHS2 HUMAN

CCGCCATGAACCCCGGCTCGATTGTCGGCCGGAACCCGCAGGAGGACTTCGAGCTGA
 TTCAGCGCATCGGCAGCGCACCTACGGCGACGCTACAAGGCACGGAAATGTTAACACTG
 GTGAATTAGCAGCAATTAAAGTAATAAAATTGGAACCAGGAGAACACTTGCAGTTGTGC
 AGCAAGAAATTATTATGATGAAAGACTGTAAACACCCAAATATTGTTGCTTATTTGGAA
 GCTATCTCAGGCAGATAAGCTTGGATTGCACTGGAGTTGAGTTGAGGGTTCTTAC
 AGGATATTATCACGTAACCTGGACCTCTGTCAAGACTGCAAATTGCATATGTTAGCAGAG
 AAACACTGCAGGGATTATATTATCTCACAGTAAAGGAAAAATGCACAGAGATATAAGG
 GAGCTAACATTCTATTACGGATAATGGCATGTGAAATTGGCTGATTTGGAGTATCTG
 CACAGATAACAGCTACAATTGCCAACGGAAAGTCTTCACTGGCACACCATATTGGATGG
 CTCCAGAAGTTGCAGCTGTTGAGAGGAAGGGGGTTACAATCAACTCTGTGATCTCTGGG
 CAGTGGGAATCACTGCCATAGAACCTGCAGAGCTTCAGCCTCTATGTTGACTTACACC
 CAATGAGAGCATTATTCATGACAAAAAGCAATTTCAGCCTCTAAACTAAAGGATA
 AAATGAAATGGTCAAATAGTTTCACTTGTGAAATGGCACTTACCAAAATCCGA
 AAAAGACCTACTGCTGAAAAATTATTACAGCATCCTTGTAAACACACATTGACAC
 GGTCTTGGCAATGAGCTGTTGGATAAGTAAATAATCCAGATCATTCCACTTACCATG

Fig. 9L
 44/76

ATTCGATGATGATGATCCTGAGCCTCTTGTGCTGTACACATAGAATTCACTCAACAA
 GTAGAAACGTGAGAGAAGAAAAAACCGCTCAGAGATAACCTTGCCAAAGTGAATTTG
 ATCCACCTTAAGAAAGGAGACAGAACCATCATGAACCTCCCACAGTGTGGTTTT
 TGGACAGTTCAGAAGAAATATACTACACTGCAAGATCTAATCTGGATCTGCAACTGGAAT
 ATGGACAAGGACACCAAGGTGGTTACTTTTAGGTGCAAACAAGAGTCTTCTCAAGTCTG
 TTGAAGAAGAATTGCATCAGCGAGGACACGTGACATTAGAAGATGATGAAGGAGATG
 ATGATGAATCTAACACTCAACTCTGAAAGCAAAAATTCCACCTCCTTGCACCAAAGC
 CTAAGTCTATCTCATACCACAGGAAATGCATTCTACTGAGGATGAAAATCAAGGAACAA
 TCAAGAGATGTCCCAGTCAAGGGAGGCCAGCAAAGCCATCCAAGTTCCACCTAGACCAC
 CACCTCCCAGATTACCCCCACACAAACCTGTTGCCTTAGGAAATGGAATGAGCTCCTTCC
 AGTTAAATGGTGAACGAGATGGCTCATTATGTCAACAACAGAATGAAACATAGAGGCACAA
 ACCTTCAAGAAAAGAAAAGATGTACCAAAGCCTATTAGTAATGGTCTTCTCCAA
 CACCTAAAGTCATATGGGTGCATGTTTCAAAAGTTTAATGGGTGTCCCTTGA
 TTCACTGTGCATCATGGATAAACCCAGATAACAAGAGATCAGTACTTGATATTGGT
 CGAAGAAGGGATTTACCCCTCAATCTTAATGAACTTCATGAAACATCAATGGAACAGC
 TATCCCTCGAAGGTGTACATGGTTGTATGTAATGAAACAATTGCTGCTATCAATATCTG
 GTAAAGCTTCTCAGCTTATTCCATAATTACCAAGGGCTTTGATTATGCAAGACAA
 TGCAAAAGTTACCTGTTGCTATTCCAGCACACAAACTCCCTGACAGAATACTGCCAAGGA
 AATTTCTGTATCAGCAAAATCCCTGAAACCAAATGGTGCAGAAGTGTGTGTAA
 GAAATCCTTACACGGGCATAAATACCTATGTGGAGCACTCAGACTAGCATTGTTCTAT
 TAGAATGGTTGAACCAATGCAGAAATTATGTTAATTAGCACATAGATTTCCTATAC
 CATGTCCACTTAGAATGTTGAAATGCTGGTAGTTCTGAACAGGAGTACCTTGT
 GTGTTGGTGTAGAGGAGACTTCAACCAAGTGGTCGATTGAGACGGTCAATC
 CAAATTCTACCTCTCATGGTTACAGAATCAGATAACCCACAGACAAATGTTACTCATG
 TAACCCAACTGGAGAGAGATACCATCCTGTATGCTGGACTGTTGTATAAAAATGTA
 ATCTCCAAGGAAGATTAAAATCTAGCAGGAAATTGTCATCAGAACTCACCTTGATTTC
 AGATTGAATCAATAGTGTGCCTACAAGACAGTGTGCTAGCTTCTGAAACATGGAATGC
 AAGGTAGAAGTTAGATCTAATGAGGTAACACAAGAAATTCAAGATAGCACAAGAATT
 TCAGGCTGCTGGATCTGACAGGGCGTGGTTGGAAAGTAGGCCAACTGATAACCCCA
 CAGCAAATAGCAATTGTACATCCTGGCGGGTCACTGAAACAGTTACTGAGAATTGTTGT
 GCTTGACAGTTAATCTAGAAAGAAACACTACCAACTGCAACATTAAATGGATGCTG
 AAGCTGTACAAAGCTGCAGTAACCTGTCTCAGTTACTTGTAAATTATTGTGGCATGA
 GATAAGATGGGAAAATTGTTAAGTGGTATGGATATTAGCATATTGAACCA
 CAAGTGCTTAATTCTATTGTAAATCTTGTACATATAAGCAGTATTCTGTGAA
 ACTTCATATTGCTGAAGACACACTAAGAATTATGAGATAATGTA
 ACTTTGAGATGTCTTATCTGTACAGATGTAATGTTGATGAAATGCAATTGGGTTA
 ATATTAAAGAATTCTTAGTATATTCTGGGTGTGGCTATATTACAAATGGGATGCTG
 GCAATGAAACAATACATTAAACACTATTGTATTATTATTATGTAATTAGTAATATGA

Fig. 9M

45/76

ATATAAATCTTGTAACTTTAAAATTGTAATGGAGGCCTGTAATCATTATAATCTTTT
 AATTAAATGCAAGTACACTGGTGTATATTGACAAAGTATTGATATGTGATGTATT
 AAGTCACAAAAGTAAGCTGTGACATTGTCTATAAGCATTGGCTCACAAATGTATTGG
 ATTGTTTCTATGTGAAGCAAACCAATTATAATTAAACCACATGTTGAGTAACGGTCTT
 TTTATATTAAAGCAGAACCTGTAAGATTGCTTGTCTTGTCTAAAAACAAATACCTTG
 ACATTTTGAAATCACAGAACAGCTTACCATGATAGAATACTGCAATTGTTGAGATT
 ACAGTATGCACAAAGAATTAAATTAGCATTATTAAAGAGTCCTCACTAAACATTGATATG
 ATCACACTGAAGAACTGTAACATTCCATAGAGTGAAGTGGTCAAATTCTCTTGGATT
 TTTACTTTGTTGCCATTATTATGATCCTTCTATTCTTTGACTTAGAGTATT
 ATACATGGCCAAAATAATTAGTTACTACCTCATACAAACAATAATGGTACTACACA
 TCACAGGAACCTAGTTGGTTAAGTCATTGATTGCTTTCCAATGGAATATGT
 ATATACAGGTTTAGCAAAATGCACACTTTGGCTCTTTGGTATATGTTCTTATAT
 TTTAATGTGAGTATACACTAAGAACAAACTAAATTGTGATTATGATCTCATTATT
 TTAATGATAATGGTTAAAATATGTTCTGATTGACATATTGTAACAAACATGTTT
 TTT

SEQ ID NO: 19 SULU1 HUMAN

GGGAGGGCTTGTGGCGCCGGCGGGCGGGCTCGTGGAGAGTGGACGCAACGCCG
 AGACCGCGAGCAGAGGCTGCGCACAGCCGGATCCGGCACTCAGCAGCGGACCCAGGAT
 CCGCCGGGAAACAAGCCACAGGAGAGCGACTCAGGAACAAGTGTGGAGAGGAAGCGCG
 GCGGCGGCCGGGCCGGGGTGGTACAGCAGGTCTGAGGTTGCATCATAAATACAAA
 GGACTGAAGTTATAAAAGAGAAAAGAGAAGTTGCTGCTAAAATGAATCTGAGCAATATG
 GAATATTGTGCCACACACAAAAAGGTACTGAAGATTACCCCCAAAAAAAATTGTCA
 ATGAGAAATAAGCTAACTGATATCAAAAGCAGAGCCTGCTCTACTGGCCATCATGCGT
 AAAGGGGTGCTGAAGGACCCAGAGATTGACGATCTATTCTACAAAGATGATCCTGAGGAA
 CTTTTATTGGTTGCATGAAATTGGACATGGAAGTTGGAGCAGTTATTGCTACA
 AATGCTCACACCAATGAGGTGGCAATTAGAAGATGCTTATAGTGGAAAGCAGACC
 CATGAGAAATGGCAAGATATTCTAAGGAAGTTAAATTTCAGACAATTGAAGCATCCT
 AATACTATTGAGTACAAAGGCTTTACTTGAAAGAACACACTGCTTGGTTGGTATGGAA
 TATTGCTTAGGCTAGCCTCTGATTAGAAGTTCTACAAAAAACCACTTCAGGAAGTG
 GAGATCGCTGCCATTACTCATGGAGCCTGCTGGACTAGCCTACCTACATTCTCATGCA
 TTGATTCTAGGGATATTAAAGCAGGAAATTCTTCTAACAGAGCCAGGTAGGTAAAA
 CTAGCTGATTGGATCTGCTTCAATGGCTCTCCTGCCACTCCTCGTGGCACACCT
 TACTGGATGGCTCCAGAGGTGATCTAGCTATGGATGAAGGACAGTATGATGGAAAGTT
 GATATTGGTCAATTGGCATCATTGTATTGAATTGGCGAACGGAAAGCCGCCCTTTC
 AACATGAATGCAATGAGTGCCTTATATCACATTGCCAGAACATGACTCCCCAACGTTACAG
 TCTAATGAATGGACAGACTCCTTAGGAGATTGTTGATTACTGCTTGCAGAAAATACCT
 CAGGAAGGCCAACATCAGCAGAACATTAGGCATGACTTGTGACGGACGCC

Fig. 9N

46/76

CTACGTGTCCTCATTGACCTCATACAGAGGACAAAGATGCAGTCGTGAGGCTAGATAAC
 CTACAGTACCGAAAAATGAAAAAAACTTTCCAAGAGACACGGAAATGGACCCCTGAAT
 GAGTCACAGGAGGATGAGGAAGACAGTGAACATGGAACCAGCCTGAACAGGGAAATGGAC
 AGCCTGGGAGCAACCATTCCATTCCAAGCATGTCGTGAGCACAGGCAGGCCAGAGCAGC
 AGTGTAAACAGCATGCAGGAAGTCATGGACGAGAGCAGTCCGAACCTGTATGATGCAC
 GATGACGAAAGCACAATCAATTCCAGCTCCTCGTGCATAAGAAAGATCATGTATT
 ACAAGGGATGAGGGGGCACGGCGATCCCAGGCCGTAGCCGCGGGCTACCCAGTCAGTT
 CAGAGCCAGGCCCTCCACTACCGGAACAGAGAGCGCTTGCCACGATCAAATCAGCATCT
 TTGGTTACACGACAGATCCATGAGCATGAGCAGGAGAACAGAGTTGCAGGGAAACAGATGTCA
 GGTTATAAGCGGATGCGGCCAGCACCAGAAGCAGCTGATGCCCTGGAGAACAGCTG
 AAGGCTGAGATGGACGAGCACCGCCTCAAGCTACAGAAGGAGGTGGAGACGCATGCCAAC
 AACTCGTCATCGAGCTGGAGAACAGCTGGCCAAGAACAGTGGCTATCATAGAAAAGGAG
 GCAAAGGTAGCTGCAGCAGATGAGAAGAACAGTCCAGCAACAGATCTTGGCCCAGCAGAAG
 AAAGATTTGACAACCTTCTAGAAAGTCAGAACAGCAGTATAAGATTTGTAAGGAAAAAA
 ATAAAAGAGGAAATGAATGAGGACCATAGCACACCCAAAGAAAGAGAACAGAGCGGATC
 TCCAAACATAAGAGAACTTGAGCAGCACACACAGGCTGAAGAGGAAGCCCACCTCTCACT
 CAACAGAGACTGTACTACGACAAAATTGTCGTTCTCAAGCGGAAAATAATGATCAAG
 CGGCACGAGGTGGAGCAGCAGAACATTGGGAGGAACATAAAAGAGGACCCAGAAG
 GAGATGGAGCATGCCATGTAATCCGGCACGAGCTCCACCCAGAGAGCTAGAGTACAGG
 CAGCTGACACGTTACAGAACAGCTACGATGGATCTGATCCGTTACAGCACCAGACGGAA
 CTGGAAAACCAGCTGGAGTACAATAAGAGGGAGAACAGACTGCACAGAAAGCATGTC
 ATGGGACTTCGGCAACAGCCAAAACTAAAGGCCATGAAATGCAAATTAAAACAG
 TTTCAGGACACTTGCAAAGTACAGACCAACAGTATAAGCACTCAAGAACATCACCAGTTG
 GAAGTTACTCCAAGAACATGAGCACAAAACATCTAAAGACACTGAAAGATGAGCAGACA
 AGAAAACTTGCCATTGGCAGAGCAGTATGAACAGAGTATAATGAAATGATGGCTCT
 CAAGCGTTACGGCTAGATGAGGCTCAAGAACAGCAGAACATGCCAGGCCCTGAGGCTACAGCTC
 CAGCAGGAAATGGAGCTGCTAACGCTTACCAAGAGCAAAATCAAGATGCAAACAGAGGCA
 CAACATGAACGTGAGCTCAGAACAGCTAGAGCAGAGAGTGTCTCGCGAGAGCACACCTT
 GAGCAGAACATTGAAGAGGGAGCTGGCTGCCCTTCAGAACGGAGCAGCAGAGAACGAGAACGAGAATAAAG
 AACCTATTGAAAGGCAAGAGCAGAGATTGAAACTTTGACATGGAGAGGCTCAGAATG
 GGATTGGAAATTGGTACATTAGATTCTAAAGAGGACTACAGATGAGATTAAATT
 TTTGCCATTACAAAAAAAGAAAACAGAAAAAAATTGAAACCTTACAGGCCCTGCAA
 AACCCACATTCCCCATTAAACGGGCCTGCTCTCACTCTCTCTCTTACTCTTACTG
 ACATCGTGTGGACTAGTGCCTGTTATTCTTACTCCATCAGGGGCCCCCTCTCCCC
 CGTGTCAACTTCAGTGCCTGGCAAAACCTGGCGTCTCTTCTATTACAGTACACGTCA
 CAGTATTGATGTGATTCAAAATGTTCAAGTAAAACCTTGGAGACAGTTAACAAAACC
 AATAAACCAACAACAAAAAGTGGATGTATATTGCTTAAGCAATCACTCATTACCACC
 AATCTGTGAAAGTAAAGCAAAAATAATAATAATGCCAAGGGGGAGAGAGACACAA

Fig. 9 O

47/76

TATCCGCAGCCTAACCTTAACAGCTGCTGCATTATTTATTTATTTATTTTG
 GTATTTATTCATCAGGAATAAAAAAAACAAAGTTTATTAAGATTGAAAATTGATACA
 TTTACAGAAACTAATTGTGATGTACATATCAGTGGTACATATTACTTTGGGG
 ACGGGGGGTGGGTGGGTGAAGAGATCTTGTGATTTAAGAACCTGCTGGCAAGAGTT
 AACCTGTCTCAGCATATTCTGATTGTATCATAATCATTCTGCTGTTGAGAGGATGT
 GAATACACTTAAGGAGCTCACAGAATCCCAGTAGCACAAATTGGGCTTGGCAAATCGT
 TATTTGTGTATAGAAGGAATTAAAGGAGAGGTATTACTTATTTCATATTGTATTTAA
 CTGTTCTCTGATCAAATTCTTACTTCTCCTGTTCTCCCCACCTCCCTCCTT
 TCCAGTTCACTTGGAGTTAACACTGTCTCAATCAGATCATCTGATCTTTCT
 TTATCTCCCTTCCCTTAAGTCCCATTCTGGTATAAATATTGATTACAC
 TTTCAAACGTGTATTTCTTACAATAAAAAATGATGAAAAAAAAAAAAAA

SEQ ID NO: 20 SULU3 HUMAN

TATTGAATTGGCGGAACGGAAGCCTCTTATTAATATGAATGCAATGAGTGCCTTATA
 TCACATAGCCAAAATGAATCCCCTACACTACAGTCTAATGAATGGTCTGATTATTTG
 CAACTTGTAGATTCTGCCTCCAGAAAATCCCTCAAGATGACCTACATCAGAGGAAC
 TTAAAGCACATATTGTTCTCGGGAGCGCCCTGAAACCGTGTAAAGATCTCATTCA
 GAGGACAAAGGATGCAGTAAGAGAGCTGGACAATCTGAGTATCGAAAGATGAAGAAAC
 CCTTTCCAGGAGGCACATAATGGACCAGCAGTAGAAGCACAGGAAGAAGAGGAAACA
 AGATCATGGTGTGGCCGGACAGGAACAGTTAATAGTGTGAAAGTAATCAATCCATTCC
 CAGCATGTCCATCAGTGCAGCAGCCAAAGCAGTAGTGTAAACAGTCTCAGATGTCTC
 AGATGACAAGAGTGAGCTAGACATGATGGAGGGAGACCACACAGTGTCTAACAGTTC
 TGTTATCCATTAAAACCAGAGGAAGAAAATTACAGAGAAGAGGGAGATCCTAGAACAAAG
 AGCATCAGATCCACAATCTCACCCCAAGTATCTGTCACAAATCACACTATGTAATCG
 AGAACACTTGCTACTATAACGGACAGCATCACTGGTTACGAGGCAAATGCAAGAACATGA
 GCAGGACTCTGAGCTTAGAGAACAAATGTCTGGCTATAAGCGAATGAGGGGACAACATCA
 AAAGCAACTGATGACTCTGGAAAACAAGCTAAAGGCTGAGATGGATGAACATGCCCTCAG
 ATTAGACAAAGATCTGAAACTCAGCTAACATTGCTGCAGAAATGGAGAAACTTAT
 CAAGAAACACCAGGCTGCCATGGAGAAAGAGGCTAAAGTGTGTCATGAAGAGAAAAA
 ATTCAGCAACATATTCAAGGCCAACAGAAGAAAGAAACTGAATAGTTTCTGAGTCCA
 GAAAAGAGAGTATAAAACTCGAAAAGAGCAGCTTAAAGAGGGAGCTAAATGAAAACCAGAG
 TACCCCCAAAAAGAAAACAGGAGTGGCTTCAAAGCAGAAGGAGAATACAGCATT
 CCAAGCAGAAGAAGCTAACCTTCTCGACGTCAAAGACAATACCTAGAGCTGGAATG
 CCGTCGCTTCAAGAGAAGAATGTTACTTGGCGTCATAACTTAGAGCAGGACCTTGTCTG
 GGAGGAGTTAACAAAAGACAGACTCAGAAGGACTTAGAGCATGCCATGCTACTCCGACA
 GCATGAATCTATGCAAGAACTGGAGTTCCGCCACCTCAACACAATTAGAAGATGCGCTG
 TGAGTTGATCAGATTACAGCATCAAACGTGACTCACTAACAGCTGGAATATAATAAGCG
 AAGAGAACGAGAACTAAGACGAAAGCATGTCATGGAAGTTGACAAACAGCCTAACAGAGTT

Fig. 9P
48/76

GAAGTCTAAAGAACTCCAAATAAAAAGCAGTTCAGGATACGTGCAAAATCCAAACAG
ACAGTACAAAGCATTAAAGAAATCACCTGCTGGAGACTACACCAAAGAGTGAGCACAAAGC
TGTTCTGAAACGGCTCAAGGAGGAACAGACCCGGAAATTAGCTATCTTGGCTGAGCAGTA
TGATCACAGCATTAAATGAAATGCTCTCCACACAAGCCCTGCCTTGGATGAAGCACAGGA
AGCAGAGTGCAGGTTTGAAGATGCAGCTGCAGCAGGAACCTGGAGCTGTTGAATGCGTA
TCAGAGCAAAATCAAGATGCAAGCTGAGGCACAACATGATCGAGAGCTTCGCAGCTTGA
ACAGAGGGTCTCCCTCCGGAGGGCACTCTAGAACAAAAGATTGAAGAAGAGATGTTGGC
TTTGCAGAATGAGCGCACAGAACGAATAGAACGCCCTGTTGGAACGTCAAGCCAGAGAGAT
TGAAGCTTTGACTCTGAAAGCATGAGACTAGGTTTAGTAATATGGTCCTTCTAATCT
CTCCCCCTGAGGCATTAGCCACAGCTACCCGGAGCTTCTGGTGGTCACACAACCTAC
TGGGGGTCCAGGACCTCACTGGGGTCACTCCCATGGGTGGCCCACCACAGCTTGGGGCCA
TCCAATGCAAGGTGGACCCAGCCATGGGTACCCCTCAGGGCCAATGCAAGGGGTAC
TCGAGGTAGCAGTATGGGAGTCCGCAATAGCCCCCAGGCTCTGAGGCGGACAGCTTCTGG
GGGACGGACGGAGCAGGCATGAGCAGAACGACGAGTGTCACTTCACAAATATCCAATGG
GTCACACATGTCTTACATAACTTAATAATTGAGAGTGGCAATTCCGCTGGAGCTGTCT
GCCAAAAGAAACTGCCTACAGACATCATCACAGCAGCCTCCTCACTTGGTACTACAGTG
TGGAGCTGAGTGCATATGGTATATTTATTCACTTTGTAAAGCGTTCTGTTTGTGTT
TACTAATTGGATGTCAAGTACTTGGCTGCCGGTTGTTGTTGGGAAATTTG
AAAAGTGGAGTTGATATTAATAATGTGTATGTGTACATATATACACACACAT
ACACATATATTATGCATGTGGTAAAAGAATTGGCTAGATAGGGATTCTGAACACT
GCAAAAATAGAACGTAGCAAAATGGCTTCAGTTATCACTTTGGTGTCTGTATCCTAAG
AAGTTTCTGAAAAGATCTAAAGCCTTTATCCCATATCCAAATTCTTATGAGCCACTC
ACAGCAGGCAGCATATGTTGAAATAAGTTATTACTGGTACACACCTGCATTGCCTCACCA
GTGTATTTATTGTTATTAATTGATCTGACTTCTCAGCCTCATTGGACTAAAAAGA
AAGCAGAAATCCATGAACACATTGCTCTCGGCCCTTGGCTAAGATCAAGTGTAGAAAT
CCATGAACACTAAAGGACTTCATTGATTTTCAGAGAGTAGAAAACAACCTAGTTTTC
TTTTTCTGAATCGTCATAGGCTTGTGAGTGTATTTGTCCATTCAATTGTGCCTCT
TTGTATTATGATAAGATGGGGTACTTAAGGAGATCACAAGTTGTGAGGATTGCATTA
ACAAACCTATGAGCCTCAATGGGAAGACCAGAACGGGTGAGAGGGGCCCTGAAAGTTCA
TATGGTGGGTATGTCCCAGCAGAGTGAGGAGATGAAGCTTACGTGTCTGACGTTTG
TTGCTTATACTGTGATATCTCATCCTAGCTAAGCTCTATAATGCCAAGACCCAAACAG
TACTTTACTTGTGTACAAAAACAAAGACATATAGCCAATACAAATCAAATGCCGA
GGTGTGATGCCATATTGCAAATTGCATCTATTGAAATTCTGTCACACTACATAGA
CATATTGTTATCTCCTTGGCTTATGTGATTTCTGTTACAAGTAGAAATGCCAATT
ATTTAAATGTTAGTGCCACAGTGAACCAAGGAGTCAGTGTGAGCCAATGACTTACCA
GCTGACTAATCTTCATCACCAGTGTAGATTTGCTGCATGTGAGGTCTTACCA
TTGCTGTTTGTGCTGCAGTACTTACAAACTTCTAGTTGAGACTTAGTGA
TTGGCATCAAGTTAACATCACACAATAGGAAACCAACTTCCACACAAGTCTCAAGCCTCA

Fig. 9Q
49/76

GTGCTAAAGTACTACTGAAAAGGAACCTAGGAAGTTGGCCAATT

SEQ ID NO: 21 SULU3 MURINE

GCAGGATGCCATCAACTAACAGAGCAGGCAGTCTAAAGGACCTGAAATTGCAGAGCTCT
 TCTTCAAAGAAGATCCGAAAAGCTCTCACAGATCTAGAGAAATCGGCCATGGGAGCT
 TTGGAGCAGTATATTTCACGAGATGTGCGTACTAATGAAGTGGTGGCCATCAAGAAAA
 TGTCTTATAGTGGAAAGCAGTCTACTGAGAAATGGCAGGATATTATTAGGAAGTCAAGT
 TTCTACAAAGAATAAAACATCCCACAGTATAGAATAACAAAGGCTGCTATTACGTGAAC
 ACACAGCATGGCTTGTAAATGGAATATTGTTAGGATCTGCTCAGATTATTAGAAGTTC
 ATAAAAAGCCATTACAAGAAGTGGAAATAGCAGCAATTACACATGGTGCTCTCCAGGGAC
 TAGCTTATTACATTCTCATACCATGATCCATAGAGATATCAAAGCAGGAAATATCCTTC
 TGACAGAACCAAGGCCAAGTGAAACTTGCTGACTTGGATCTGCTCCATGGCTTCCCTG
 CCAATTCTTTGGAACACCATATTGGATGGCCCCAGAAGTAATTAGCCATGGATG
 AAGGACAGTATGATGGCAAAGTTGATGTATGGTCTCTTGGAAATAACGTGTATTGAATTAG
 CCGAGAGGAAGCCTCTTATTAAATATGAATGCAATGAGTGCTTATATCACATAGCCC
 AAAATGAATCCCTACACTACAATCTAATATGAATGATTCTGCCTCCAGAAAATCCCTC
 AAGATGCCCTACATCAGAGGAACCTTAAAGCACATGTTGTTCTCGAGAGCGCCCTG
 AACACAGTGTAAATAGATCTTATTCAAAGGACAAAGGATGCAGTAAGAGAGCTGGACAATC
 TGCAGTATCGAAAGATGAAGAAACTCCTTCCAGGAGGCACATAATGGGCCAGCGGTAG
 AAGCACAGGAAGAAGAGGGAGCAAGATCATGGTGTGGCGAACAGGAACAGTGAATA
 GTGTTGGAAGCAATCAGTCTATCCCTAGTATGTCTATCAGTGCAGCAGTCAAAGCAGCA
 GTGTTAAATAGTCTTCCAGATGCATCAGATGACAAGAGTGAGCTAGACATGATGGAGGGAG
 ACCATACAGTGTCTAACAGTTCTGCATCCACTTAAACCTGAGGAGGAAAATTACC
 AGGAAGAAGGAGATCCTAGAACAAAGAGCATCAGACCCACAGTCTCCCCCTCAGGTGTCTC
 GTCACAAGTCACATTATCGTAATAGAGAACACTTGCAACCATAACGAAACAGCATCACTGG
 TTACAAGACAGATGCAAGAACATGAGCAGGACTCTGAACCTAGAGAACAGATGTCTGGTT
 ATAAGCGGATGAGGCACAGCATCAAAGCAGCTGATGACGCTGGAAAATAACTGAAGG
 CAGAGATGGACGAACATGGCTCAGATTAGACAAAGATCTGAAACTCAGCGTAACAATT
 TCGCTGCAGAAATGGAGAAACTTATTAGAAACACCAAGCTGCTATGGAAAAAGAGGCTA
 AAGTGTGGCCAATGAGGAGAAAAATTCCAGCAACACATTAGGCTCAACAGAAAAAG
 AACTGAATAGCTTTGGAGTCTAAAAAGAGAAATATAACTCGCAAAGAGCAGCTTA
 AGGAGGAGCTGAATGAAAACCAGAGCACACCTAAAAAGAAAAGCAGGAATGGCTTCAA
 AGCAGAAGGAGAATATACGATTTCAGGCAGAAGAAGCTAATCTCTCGACGTC
 AAAGGCAGTATCTAGAGCTAGAATGTCGTCGCTTCAAAGAAGAATGTTACTTGGCGAC
 ATAACCTGGAACAGGACCTTGTCAAGGAGGAGTTAAACAAAAGCAGACTCAAAGGACT
 TGGAACATGCAATGCTATTGCACAGCATGAATCAATGCAAGAACTGGAGTTCGCCATC
 TCAACACTATTAGAAGATGCGCTGTGAGTTGATCAGACTGCAGCATCAAACGTGAGCTCA
 CTAACCAGCTAGAGTACAATAAGAGAAGGAACTGGAACTGAGGCAGAACATGTCATGG

Fig. 9R
 50/76

AAGTTGACAACACCTAACAGACTGAAGTCTAAAGAACTCCAATAAAAAGCAGTTCA
AGGATACCTGCAAATTCAAACCAAGACAGTACAAAGCATTAAGGAATCACCTACTGGAGA
CTACACCAAAGAACATGAGCACAAAGCAATC

SEQ ID NO: 25 GEK2 HUMAN

CGAAGCCACAGCCCGAGCCCGAGCCCGAGCCCGAGCCGGCGCCACCGCGCCCCGGCCAT
GGCTTTGCCAATTCCGCCGCATCCTGCGCTGTCTACCTTGAGAAGAGAAAGTCCCG
CGAATATGAGCACGTCCGCCGCACCTGGACCCCAACGAGGTGTGGAGATCGTGGCGA
GCTGGCGACGGCGCTTCGGCAAGGTTACAAGGCCAAGAATAAGGAGACGGGTGCTTT
GGCTGCGGCCAAAGTCATTGAAACCAAGAGTGAGGGAGCTGGAGGACTACATCGTGA
GATTGAGATCCTGCCACCTGCACCCCTACATTGTGAAGCTCCTGGAGCCTACTA
TCACGACGGGAAGCTGTGGATCATGATTGAGTTCTGTCCAGGGGGAGCCGTGGACGCCAT
CATGCTGGAGCTGGACAGAGGCCACGGAGCCCAGATAACAGGTGGTTGCCGCCAGAT
GCTAGAACCCCTCAACTCCTGCACAGCAAGAGGATCATCCACCGAGATCTGAAAGCTGG
AACGTGCTGATGACCCCTGAGGGAGACATCAGGCTGGCTGACTTGGTGTCTGCCAA
GAATCTGAAGACTCTACAGAAACGAGATTCTTCATCGGCACGCCCTACTGGATGGCCCC
CGAGGTGGTCATGTGTGAGACCATGAAAGACACGCCCTACGACTACAAAGCCGACATCTG
GTCCTGGCATCACGCTGATTGAGATGGCCAGATCGAGCCGCCACACCACGAGCTCAA
CCCCATGCCGGTCTGCTAAAGATGCCAAGTCGGACCCCTCCACGCTGCTCACGCCCTC
CAAGTGGCTGTAGAGTCCGTGACTTCTGAAGATAGCCCTGGATAAGAACCCAGAAC
CCGACCCAGTGCCCGCAGCTGCTGGAGCATCCCTCGTCAGCAGCATCACAGTAACAA
GGCTCTGCCGGAGCTGGTGGCTGAGGCCAAGGCCAGGTGATGGAAGAGATCGAAGACGG
CCGGGATGAGGGGGAGAGGGAGGACGCCGTGGATGCCGCCACCCCTGGAGAACCATAC
TCAGAACTCCTCTGAGGTGAGTCCGCCAAGCCTCAATGCTGACAAGCCTCTGAGGGAGTC
ACCTCCACCCGCTGGACCCAGCCAGTCTCAGGACAGTGTGAATGAGCCCTGAGCCA
GCCCTCTGGGACAGATCCCTCAAACACCAGTCCCCAGTCGTGGCCCTGGAAATGA
GAACGCCCTGGCAGTGCCTGTGCCCCCTGCCAGTCCCACCCGTGTCATGGATGCCAG
AATTAGGTAGCCAGGAGAACAGCAAGTTGCTGAGCAGGGTGGGACCTCAGCCAGCAGC
AACAGATCTCAAAGGCCAGCCAGGCCAGCAGCAGCAGCAGCAGCTGGAGACCTTGGG
TGGGGAGAACCTGGCAATGGCAGCCTGGAGCCACCTGCCAGGCAGCTCCAGGCCCTTC
CAAGAGGGACTGGACTGCAGCAGCCTCTGCACCTCTGAGAGCATGGACTATGGTACCAA
TCTCTCCACTGACCTGTCGTGAAACAAAGAGATGGGCTCTGTCCATCAAGGACCCGAA
ACTGTACAAAAAAACCTCAAGCGGACACGCAAATTGTGGTGGATGGTGTGGAGGGTAG
CATCACCAACCTCCAAGATCATCAGCGAAGATGAGAAGAAGGATGAGGAGATGAGATTCT
CAGGCCAGGAACCTGGAGAGCTTCGGCTGCTCCAGAAAGAACAGCATCGGAACCCAGAC
CCAGCTGAGTAACAAGCATGAGCTGCAGCTGGAGCAAATGCATAAACGTTTGACAGGA
AATCAACGCCAAGAACAGAGTTCTTGACACGGAATTAGAGAACCTGGAGCGTCAGCAAA
GCAGCAAGTGGAGAACATGGAGCAAGACCATGCCGTGCCGCCGGAGGGAGGCCAGGCG

Fig. 9S
51/76

GATCCGCCTGGAGCAGGATCGGGACTACACCAGGTTCCAAGAGCAGCTAAACTGATGAA
 GAAAGAGGTGAAGAACGAGGTGGAGAAGCTCCCCGACAGCAGCAGGAAAGGAAAGCATGAA
 GCAGAAGATGGAGGAGCACACGAGAAAAGCAGCTTCTTGACCAGGGACTTGTAGCCAA
 GCAGAAGGAGGACCTGGAGCTGGCATGAAGAGGCTACCACCGACAACAGGCAGGAGAT
 CTGTGACAAGGAGCGCGAGTGCCTCATGAAGAAGCAGGAGCTCCTCGAGACCAGGGAGC
 AGCCCTGTGGAGATGGAAGAGCACCAGCTGCAGGAGAGGCACCAGCTGGTGAAGCAGCA
 GCTCAAAGACCAGTACTTCCTCCAGCGGACAGCTGCTGCGCAAGCATGAGAAGGGAGCG
 GGAGCAGATGCAGCGCTACAACCAGCGCATGATAGAGCAGCTGAAGGTGCGGAGCAACA
 GGAAAAGGCAGGGCTGCCAAGATCCAGAGGAGTGAAGGGCAAGACGCGCATGCCATGTA
 CAAGAAGAGCCTCACATCAACGGCGGGGAGCGCAGCTGAGCAGCGTGAAGAAGATCAA
 GCAGTTCTCCCAGCAGGAGGAAGAGGAGAAGTCGGAGCGGCTGCAGAACAGCAGAA
 ACACGAGAACCAAGATGCGGGACATGCTGGCGAGTGCAGAGCAACATGAGCGAGCTGCA
 GCAGCTGAGAATGAAAAGTGCCACCTCCTGGTAGAGCACGAAACCCAGAAACTGAAGGC
 CCTGGATGAGAGCCATAACCAGAACCTGAAGGAAT

SEQ ID NO: 27 PAK4 HUMAN

CGTTCCCTGGCTTCCGCTCCGAGGCCTGCGGAGGACTGGCCAGCAAGGTCCAGGTC
 TTCCCTCTCCTTAGCGCTAAGAGAGAGGCCCAGTGCAGGGTGAAGGAGTCGCAGGAAGAG
 GCGGAAGGCAGGGAAAGGCACCATGTTCCGCAAGAAAAAGAAGAAACGCCCTGAGATCTC
 AGCGCCACAGAACTTCCAGCACCGTGTCCACACCTCCTCGACCCAAAGAAGGCAAGTT
 TGTGGGCCTCCCCCACAATGGCAGAACATCCTGGACACACTGCGGCGCCCAAGGCCGT
 GGTGGACCCCTCGGAATCACACGGGTGCAGCTCCAGCCATGAAGACAGTGGTGGGG
 CAGCGCGATGCCTGTGGATGGTACATCTCGGGCTGCTAACGACATCCAGAAGTTGTC
 AGTCATCAGCTCCAACACCCCTCGTGGCCAGCCCCACCGAGCCGGCGGGCACAGTC
 CCTGGGGCTGCTGGGGATGAGCACTGGGCACCGACCCAGACATGTACCTCCAGAGCCC
 CCAGTCTGAGCGCACTGACCCCCACGGCCTTACCTCAGCTGCAACGGGGCACACCAGC
 AGGCCACAAGCAGATGCCGTGGCCGAGCCACAGAGCCCACGGTCTGCCAATGGCT
 GGCTGCAAAGGCACAGTCCCTGGCCCCGCCGAGTTCAAGGGTGCCTCGCAGCGCTGTCT
 GCAGCTGGGTGCCTGCCTGCAGAGCTCCCCACCGAGGAGCCTCGCCCCCACGGCACCAA
 TAGGCATGGAATGAAGGCTGCCAAGCATGGCTCTGAGGAGGGCCACAGTCCCTGCCT
 GGTGGGCTCAGCCACAGGCAGGCCAGGTGGGAAGGCAGCCCTAGCCCTAAGACCCGGGA
 GAGCAGCCTGAAGCGCAGGCTATTCCGAAGCATGTTCTGTCCACTGCTGCCACAGCCC
 TCCAAGCAGCAGCAAGCCAGGCCCTCACACAGAGCAAGCCAACTCCTCTTCCGACC
 GCCGCAGAAAGACAACCCCCCAAGCCTGGGCCAGGCCAGTCCTGCCCTCGGACCA
 GCCGGTGGGGACCTTCAGCCCTTGACCACTTCGGATAACAGCAGCCCCAGAAGTCCCT
 CGGCACAGCCCCGGCCACAGGCCAGCTCCAGGCCGGTCTTCCCCAGCAGGGATCCCCCG
 CACCTGGCACGCCAGATCAGCACAGCAACCTGTACCTGCCAGGACCCACGGTTGC
 CAAGGGTGCCTGGCTGGTAGGGACACAGGTGTTGACACATGAGCAGTTCAAGGCTGC

Fig. 9T
 52/76

GCTCAGGATGGTGGACCGAGGTGACCCCCGGCTGCTGGACAGCTACGTGAAGAT
 TGGCGAGGGCTCCACCGGCATCGTCTGCTTGGCCCGGGAGAAGCACTCGGGCCGCCAGGT
 GGCGTCAAGATGATGGACCTCAGGAAGCAGCAGCGCAGGGAGCTGCTCTCAACGAGGT
 GGTGATCATGCGGGACTACCAGCACTCAACGTGGAGATGTACAAGAGCTACCTGGT
 GGGCGAGGAGCTGTGGGTGCTCATGGAGTCCTGCAGGGAGGCCCTCACAGACATCGT
 CTCCCAAGTCAGGCTGAATGAGGAGCAGATTGCCACTGTGTGAGGCTGTGCTGCAGGC
 CCTGGCCTACCTGCATGCTCAGGGTGTATCCACCGGGACATCAAGAGTGACTCCATCCT
 GCTGACCCCTCGATGGCAGGGTAAGGCTCTCGGACTTCGGATTCTGTGCTCAGATCAGCAA
 AGACGTCCCTAACGAGGAAGTCCCTGGTGGAAACCCCTACTGGATGGCTCCTGAAGTGAT
 CTCCAGGTCTTGATGCCACTGAGGTGGATATCTGGTCTCTGGGATCATGGTATTGA
 GATGGTAGATGGGGAGCCACCGTACTTCAGTGACTCCCCAGTGCAAGCCATGAAGAGGCT
 CGGGGACAGCCCCCCCACCCAAGCTGAAAAACTCTCACAGGTCTCCCCAGTGCTGCAGA
 CTTCTGGAGCGGATGCTGGTGCAGGGACCCCCAAGAGAGAGGCCACAGCCCAGGAGCTCCT
 AGACCAACCCCTTCTGCTGCAGACAGGGTACCTGAGTGCCCTGGTCCCCCTGATCCAGCT
 CTACCGAAAGCAGACCTCCACCTGCTGAGCCACCCCAAGTATGCCTGCCACCTACGCC
 ACAGGCAGGGCACACTGGCAGCCAGCCTGCCGGCAGGACTTGCCCTGCCCTCCTCTCA
 GTATTCTCTCCAAAGATTGAAATGTGAAGCCCCAGCCCCACCCCTTGCCCTTCAAGCCTAC
 TGGGCCAGGCCGGACCTGCCCTCAGTGTCTCTCCCTCCGAGTCCCCAGATGGAGACC
 CCTTCTACAGGATGACCCCTTGATATTGCACAGGGATATTCTAAGAAACGCAGAGGC
 CAGCGTTCTGGCCTCTGCAAGCCAACACAGTAGAAAAGGCTGCTGTGGTTTTAAAGGC
 AGTTGTCCACTAGTGTCTAGGCCACTGCAGAGGGCAGACTGCTGGTCTCACAGATA
 TGCTGTTCTCAGCTCCAGCTTCAACCTCGAGTCTCGAGAGGGCACGGGGTGGTTTTA
 TGACCGGAATCCGCTTCCCTCACGTCTGATGTCTGAAGGTGCAGTCCCACCTGTA
 CAGCCCCCTCCCCGCCAAGAACTGTGAATGGCCTGCTCCAGGCCATGGCTGGGGGAGGG
 GTGAGGGGACAATTCTGAGTGAAAGAGAAAGAATGGGTCGGTGGTAAGGTGCTCTCA
 CTTACAGAAATGGAGAGAACATCGTGTGTGTGTGTGTGTGTGTGTGTGTGT
 TGT
 CCTCCAGGTCAACACAGCCAGTTCAAGGAAGGCTGCCCTCTCTCCACTAACGTTCTGG
 CCTGAAGGGACCTGCTTCTGGCCTGGCTTCCACCTCTCCACTCCTGTGTCTACCTGGC
 CAGTGGAGTGGCCATGCTAACACTCCTGGAGCTCAGGAGGCTCTGAGCTTC
 TCCTGTACTGTGCATCGTGGCCAGAGACAGGAATGTAAGGATTGGCAACTGTGTTAC
 CTTCAAGTTATCTCAATAACCAGGTACAGGGACCCATTGTTCTCTCAGAACCTA
 TCTGGGAGAGAAGGCGAACCACTCCGGGTTCCATCATGTCAGGTACAGGCATCCAT
 GTGTGCAAACCATCTGCCCTAGCTGCTCCACAGACTGCTGTCTCTGTCTCCCTCGGC
 CCTGCCCTACTCAGGGCTGCTGTGAGATGGAATTCCAGGAAAGAACTTCAGGTGTCTGG
 ACCCTTCTATCTAGATAATATTTAGATTCTCTGCTCCCTAGTGACCTACCTGGGGG
 CAAAGAAATTGCAAGGACTTTTTAAGGGTCAGAGTTCAAAACAAAAGCATCTTCC
 CTAGAAATTTGTGAATTGTTGCACTTGTGCCTGTTAAATTAAATTGAGTGTCAA

Fig. 9U

53/76

AGCC

SEQ ID NO: 28 PAK5 HUMAN

GGCCAGTGGGGCGAAACTGGCAGCTGGCCGGCCCTTAACACCTACCCGAGGGCTGACAC
GGACCACCCATCCC GG GTGCCAAGGGGAGCCTCATGACGTGGCCCTAACGGGCCATC
AGCGGGGGGCTGGC ATCCCCAGTCCTCCTCCTCCGGCCTCCACCCGAGC
CCGAGGTGCCCCCAGCCCTGGAGTGCTGGACCCACGCCTCAGAGCCCCAGCTGGCCCC
TCCAGCCTGCACCCCGCCGCCCTGCTGTTCTGGGCCCCCTGGCCCCCGCTACCCACA
GCGGGAGCCACAGCGAGTATCCCATGAGCAGTTCCGGCTGCCCTGCAGCTGGTGGTGA
CCCAGGC GACCCCGCT CCTACCTGGACA ACTT CATCAAGATTGGCGAGGGCTCCACGGG
CATCGTGTGCATGCCACCGTGCAGCTCGGCAAGCTGGTGGCGTCAAGAAGATGGA
CCTGCAGCAGCAGCAGAGGGCGCAGCTGCTCTTCAACGAGGTGGTAATCATGAGGGACTA
CCAGCACGAGAATGTGGTGGAGATGTACAACAGCTACCTGGTGGGGACGAGCTGGT
GGTCA TGGAGTTCTGGAAGGAGGGCGCCCTACCGACATCGTACCCACACAGGATGAA
CGAGGAGCAGATCGCGGCCGTG CAGTGCAGGCCCTGCGGTGCTCCACGC
CCAGGGCGTCA TCCACC CGG GACATCAAGAGCGACTCGATCCTGCTGACCCATGATGGCAG
GGTGAAGCTGTCA GACTTGGTTCTGCGCCAGGTGAGCAAGGAAGTGGCCGAAGGAA
GTCGCTGGTGGCACGCCCTACTGGATGGCCCAGAGCTCATCTCCGCCCTCCACGG
GCCAGAGGTAGACATCTGGTCGCTGGGATAATGGT GATTGAGATGGTGGACGGAGAGCC
CCCCTACTTCAACGAGCCACCCCTCAAAGCCATGAAGATGATTGGGACAACCTGCCACC
CCGACTGAAGAACCTGCACAAGGTGTCGCCATCCCTGAAGGGCTCCTGGACCGCCTGCT
GGTGCAGGACCCCTGCCAGCGGGCACGGCAGCCAGCTGCTGAAGCACCCATTCTGGC
CAAGGCAGGGCCGCCCTGCCAGCATCGTCCCTCATGCGCCAGAACCGCACCAGATGAGG
CCCAGCGCCCTTCCCTCAACCAAAGAGCCCCCCC GGTCACCCCGCCCACTGAGGCC
AGTAGGGGGCCAGGCCTCCACTCCCTCCAGCCGGAGATGCTCCCGTGGCACCCACCC
TCCTTGCTGGGGTAGATGAGACCCCTACTACTGAACCTCAGTTGATCTGTGACTTT
AGAAAAAACACAGGGACTCGTGGGAGCAAGCGAGGCTCCAGGACCCCCACCCCTGGGAC
AGGCCCTCCCCCATGTTCTCTGTCTCCAGGAAGGGCAGCGGCCCTCCACACTGGAAG
TCTGCAGTGGGGTCGCTGGGGTGGAGAGAACACTAAGAGGTGAACATGTATGAGTGT
TGCACCGCGTGTGAGTGTGCATGTGTGTGTGCAAAGGTCCAGCCACCCGTCCTCCA
GCCCGCAAGGGGTGTCTGGCGCCCTGCCTGACACCCAGCCCCCTCTCCCCGTAGCCCTGCGCCCCCTT
CTGTGGCTGGATGGGAGACAGGTCAAGGGCCCCCACCCCTCCAGCCCTGCAGCAAAT
GACTACTGCACCTGGACAGCCTCTCTTTCTAGAAGTCTATTATATTGTCA TTTATA
ACACTCTAGCCCTGCCCTATTGGGGACAGATGGTCCCTGTCCCTGCGGGGTGGCCCTG
GCAGAACCAACTGCCTGAAGAACCAAGGTTCTGCCGGTCA GCAGCCCCAGCCCGCCCA
CCCCTGCCCTGAGTTAGTTACAATTAAAACATTGTCTTGTAAAAA
AAAAAAAAAA

Fig. 9V
54/76

>STLK5_h
MSSFLPEGGCYELLTIVKGFEIDLMTVNLLARYKPTGEYVTVRRIINLEACSNEMVTFLQGELHVSKLFNHPNIVPYRATFI
ADNELWVVTSFMAYGSAKDLICHTFMDGMNELAIAYILQGVLKALDYIHMGYVHRSVKAASHILISVDGKVYLSGLRSNL
SMISHGQRQRVVFHDFFPKYSVKVLPSPEVLQQNLDQGYDAKSDIYSVGITACELANGHVPFKDMPATQMLLEKLNGTVPC
LLDTSTIPAAELTMSPSRSVANSGLSDSLLTSPRPSNGDSSPSSHYPYHRTFSPHFFFVEQCLQRNPDARPSASTLNLHSF
EKQIKRRASEALPELLRPVTPTINFEGSQSQDHSGIFGLVTNLLELEVDDWEF

>STLK6_h
MSLLDCFCCTSRTQVESLRPEKQSETSIHQYLVDDEPTLSWSRPSTRASEVL
PTGTLVTKITNLENCNEERLKLQAVILSHFFFRRHPNITTYYWTVFTVGS
IRNILFGAVVRGLNYLHQNGCIHRSIKASHILISGDGLVTLSGLSHLHS
55 DLHGYNVKSDIYSVGITACELASQVPFQDMHRTQMLLQKLKGPPSPLD
GTHTVNSDRLHTPSSKTFSPAFFSLVQLCLQQDPKPSISLPPVLP
WTEPECDFPDEKDSYWEF

>STLK7_h
NRDDYELQEVIIGSGATAVVQAAycopKKEKVAIKRINLEKQTSMDELLKEIQAMSQCHHPNIVSYYTTSFVVKDELWLVM
KLSGGGSVLDIIKHIIVAKGEHK

>ZC4_h
 MAGPGGGWRDREVTDLGHPDPTGIFSLDKTIGLGGTYGRIYLGLHEKTGAFTAVKMNARKDEEDLRTELNLLRKYSFHK
 NIVSFYGFYAFFKLSPPGQRHQQLWWMCAAGSVTDVRMTSNQSLKEDDWIAYICREILQGLAHLHAHRVIHRSYD
 LTHNAEVKLVDFGVSAQVSRTNGRNSFIGTPYWMAPEVIDCDEDPRRSYDYSVGIITAEMAEAPPCLCNLQPLE
 ALFVILRESAPTVKSSGWSRKFHNFMEKCTIKNFLFRPTSANMLQHPFVRDIKNERRHVVESLTRHLTGIIKKRQK
 EKKSKYSTIRDALAKBLSPKRERAKSSEWRPEKLELSDEARRQRQRWEDIFNQHEELRQVDKDKEDESSDND
 EYFHS

Fig. 10A

IQAEVQIEPIKPYISNPKKIEVQERSPSVPNNQDHAAHHVKFSSSVQRSLLEQAQKPIDIRQRSSQNRQNWLAAASGDSKH
 KILAGKTQSYCLTIYISEVKKEEFQEGMNQKCQGAQVGLGPEGHCIWQLGESSEESPVTPRRSQSSPPYSTIDQKL
 DIHVPDGFVKVGKISPPVYLTNNEWGYNALSEIFRNDWLTPAPVIQPPPEEDGYVELYDASADTDGDDDESNTFEDTYD
 HANGNDLDDNQVQDQANDVCKDHDDNNKFVDDVNNYYEAPS
 C PRASYGRDGSCCKQDGYDGSRGKEEAYRGYGSHTANRS
 HGGSAASEDNAAIGDQEEHAANIGSERGSEGDGGKGVVRTSEESGALGLNGEENCSETDGPGLKRPA
 SQDFYQLEPLNLLITISG
 GGNEASNAIDS
 GAAPSAPDHESDNKDISSESSTQSDFSANHSSPSKGMSADANFA
 SAILYAGFVEVPEESPQKQPSEVNV
 NPLYVSPACKPLIHMYEKEFTSEICCGSLWGWNLLLGTRSNLYLM
 DRSGKADITKLIRRRPFRQIQVLEPLNLLITISG
 HKNRLRVRVYHTLWRNKILNNDPESKRQEEMLKTEACKA
 IDKLTGCEHFSQLQHEETTYIAIALKSSIHLYAWAPKSFD
 ESTAIKVFP
 TLDHKPVTVDLAIGSEKRLKIFSSADGYHLIDA
 ESEWMSDVTLPKNPLEIIIPQNIILPDCLGIGMMLT
 FNAEALSVEANEQLFKKILEMWKDIPSSIAFFECTQRTTG
 GWGQKAIEVRSLSQSRVLESELKRRSIKKLRFLCTR
 GDKLFFT
 STLRNHHSR VYFMTLGKLEELQS NYDV

56/76
 >PAKS_h

MFGKRKKRVEISAPSNF
 EHRVHTGFDQHEQKFTGLPRQWQSLIEESARRPKPLVDPACITSI
 QPGAPKTIIVRGSKGAKDG
 ALTL
 LDEFENMSVTRNSNLRDSSPPARARQENGMP
 PEETPATTARGPGKA
 GSRRGFRAGHSEAGGGSGDRRRAGPEKRP
 KSSREGSGGPQESSRDKRPLSGPDVGTPQPA
 GLASGAKLAAGRPFNTYPRADTDHP
 SRGAQGEPHDVAPNGPSAGGLAIP
 QSSSSSSRPTARGAPSPGVLGPHASEPQLAPPACTPA
 AAVPGPPGPRSPQREPQ
 RVSHEQFRAALQLVVDPGDPRSY
 LDNFIKIGESTGIVCIATVRS
 SSGKLVAKKMDLRKQQRRELLFNEVV
 VIMRDYQHENVV
 EMYNSYLV
 GDELWVVMFLEG
 GALTDIVTHTRMNEEQIAAVCLAVL
 QALSVLHAQGV
 IHRDIKSDSILLTHDGRVKLSDF
 GFCAQVSKEV
 PRRKSLV
 GTPY
 WMAPELISRLPYGPEVDIWSL
 GIMVEMV
 DGEPPYFNEPPLKAMK
 MIRDNLPPR
 LKVLVRDPAQR
 ATAELLKHPFLAKAGPPASIVPLMRQRNTR

>GEK2_h

MAFANFRRRILRLSTFEKRKSREYEHVRRLDPNEVWEIVGELGDGA
 FGKVYKAKN
 KETGALA
 AAKVIETK
 SEE
 ELEDYIV
 EIEILATCDHPYIVKLLGAYYHDGKL
 WIMIEFC
 PGGA
 VDAIMLELD
 RGLTEP
 QIQQV
 VCRQM
 LEALN
 FLHS
 KRI
 IIH
 RDLKA

Fig. 10B

GNVLMTLEGDIRLADFFGVSAKNLKTLQKRDSFIGTPYWWMAPEVVMCETMKDTPYDYKADIWSLGITLIEAMAQIEPPHHEL
 NPMRVLLKIAKSDPPTLLTPSKWSVEFRDFLKIALDKNPETRPSAAQLLEHPFVSSITSNKALRELVAEAKAEVMEETIED
 GRDEGEEDAVDAASTTLENHTQNSSEVSPPSLNADKPLEESPSTPLAPSQSQQDSVNÉPCSQPSGDRSLQTTSPPVVAPGN
 ENGLAVPVPLRKSRPVSMNDARIQVAEQQGDLSPAANRSQKASOSRPNSSALETLGGEKLANGSLEPPAQAAQGP
 SKRDSDCSSLCTSESMWDYGTNLSTDLSLNKEENGSLSIKDPKLYKKTLKTRKFVVDGEVSITTSKIISEDEKKDEEMRF
 LRRQELRELRLQKEEHRNQTQLSNKHELQLEQMHHKRFQEINAKKKFFDTENLERQQKQQVEKMEQDHAVRRREEAR
 RIRLEQDRDYTRFQEQLKLMKKKEVKNEVEKLPRQRKESWKQKMEEEHTQKKQLLDRDFVAKQKEDLEAMKRLTTDNRRE
 ICDKERECLMKKQELLRDREAALWEMEEHQLQERHQVLVKQQLKDQYFLQRHELLRKHEKEREQMQRYNQRMIEQLKVRRQQ
 QEKAIRLPKIQRSSEGKTRMAMYKKSLHINGGGSAAEQREKIKQFSQQEEKRQKSERLQQQQKHENQMRDMLAQCESNMSEL
 QQLQNEKCHLLVEHETQKLKALEEDLNQKKREQEMFFKLSEEAECPNPSTPSKAACKFF
 PYSSGDAS

Fig. 10C

58/76

Fig. 11A

ACCTTAAGTCTTTCCACGGTTATGTGCTCATTCCTTCCACCAAGAATCCATCCTTAGCGGCCAGCTG
CCCCCTGGTGCTTCTCAAGGGCATCAGTGTCTTGCCTAGCTTAACTGCTTAAAGTCCCTAAGTGTGTTAGTTCTGTTGT
TCAAGAACGGCTG

>STLK6_h
AAGGAAGATAAAACAAAAGCCTTCTTGGAAATAGATGGATTTCCTGTGTGAAACTAAAGTGAATTCAATGTCT
CTTTGGATTCTGCACCTCAAGAACACAAGTGAATCAACTGAAACCTGAAAGTCTGAAACCTGAAACTCACTCCATCA
AACTTGGTTGATGCCAACCCCCCTTCCGTACGTCATCCACTAGGCCAGTGAAGTACTATGTTCCACCAACGTTT
CTCACTATGAGCTCAAGTGAAGTAGGAAGGATTGACAACCTTGACCTTGGAAACTTGAAGGATTGAGGATTT
GGAAACACTGGTAACCTATAAAATACAAATCTGGAAAACCTGCAATGAAAGAACGGCTTGAAGAAGCTTACAGAAAGCCGTGAT
TCTATCCCACCTTTCCGGCATCCCATAACTTAACTTACAACCTATTGACAGTTTCACTGTTGAGCTGGCTTGGTTATT
CTCCATTATGGCCTATGGCTTCAACTCTTGAAGGACCTATTGAGGACTTCTCTGAAGGAATGAGTGAAGCTTAAAGA
AACATTCTCTTGGAGGGTTGAAGAGGGTCAACCAAAATGGCTGTATTCAAGGAGTATTAAAGCCAGCCA
TATCCTCATTTCTGGTATGGCTTCAAGCCATCAGTGCAGCCGTGAGCTGGCTGAGTCCAGAAACTACTGAGACAGGATT
ATAGGGCTGTGTATGATTCCCACAGTTCAAGCAGTGGATTACAGCATGTGAATTAGCCAGTGGATTACAGTGT
CATGGGTATAATGTGAAGTCAGATTTACAGTGTGGATTACAGTGTGGATTACAGTGTGGATTACAGTGTGGATT
GGACATGCCATAGAAACTCAGATGCTGTTACAGAAACTGAAAGGTCCCTATAGCCATTGGGAAAGTGTGCTTCCAGTGGAACT
AATCAGAATCCAGAATGAAAAATCCCAGTCAGGTAGCTGGACTCTGGGATTTCAAGGAACTTCAAAACCT
CACACAGTAAATAGTGAACCGATTACACACCATCCTCAAAACCTTCTGGCTTCAAGGAACTTCAAAACGCTTCA
TTTGCAACAAAGATCCGTGAAAGAAGGCCATCAGCAAGCAGTCAAGCAGTGGATTCTGTTCAAAACAGTGA
GCCAGGATTCAATACTTCACTGTTGCTTCTGCTTCAAAAGCCATCAATCATTGGCTTCAAAAGGCAATTCTAGGT
GAGCCAGAAATGTGATTTCCTGATGAAAAAGACTCATACTGGGAATTCTAGGTACAAATTACAGGAAATTATAC
CTTGACACTTCTCCCTGGCTTTCTGTTCAAAAGGGCACCAAGTTGTAAGTCCCTCTGTTCAAAAGGGCAC
GCACCTGGGAATCTTAAACCAACTCTGTTCAAAACCTTAACTTAAACCAACTCTGTTCAAAAGGGCACCAAGTTG
ATGACAAGGAAACATCAGAATTACTAACTTAGCTAGTGTCAATTATTCTGGAAATTTTCTJAAGCTGTGACTA
ACTCTT

TTTATCTCAATAATTGGCCAGTTAATTTCAGTATTTCAGGAAATGGCCCTTGCTGCCCCCTCAGAGGAC
AATGCTTCCAAGTACATCTTCTCCAGATTCTGGCTTAACTGCTATTGAGCTATTGTAAACAGGCTAGTTATCTT
ACATCAGACCCTTCTGTAGAGGGAAATGTTTCTGTTAACACTTATGGTAACACTTAAC
TGAGCCTCACTCACCTAAATGATTCACTTGAATAATACAGAAATTGTAATTGCTTTTAAAGGGGGCTAA
AGCTAAATTCTAAATTGATAATTAGCTTACAGGCCACCCGTGGGAGCTCAATAAGATTACTGAATT
AGCTAAACCTTCTACITATGTAATTAGCTTACAGGCCACCCGTGGGAGCTCAATAAGATTACTGAATT

५

60/76

SUBSTITUTE SHEET (RULE 26)

Fig. 11

61/76

CATAAGAACAGACTTGGGTGATCATCTGACCTGGTTGAGGAACAAGATAATTGATTTGAACTTCAAGATTTGATAAGCTATTGATAACAGGCTGTAACACTTCAGTGTCTCC
GCAAGAACAAATGCTGAAGACAGGAAAGCCTGAAAGCTTAAACGGCTTAAACGGTAAACAGGCTATTGCTTTGAAATCATCAATTCACCTTATGCATGGCACCAAGTGCCTTGT
AACATGAAGAACACATATAATTGCAATTGCTTGTAAACATGATTCACCTGCTTAAACACTTGATCATAAGCCAGTGAAGTTGACCTGCTTAAAGTATTTCAAAAG
GAAAGGCACTGCTTAAAGTATTTCAAAACTTCAACTTAAAGTATTTCAAAACTTCAACTTCAACTTCAACTTCAACTTCAACTTCAACTTCAACTTCAACTTCAACTTCA
ACTAAAGATTTCTCAGCAGATGGATACTGATGGAATCTGATGCTTGGGAATTGGCATGATGCTCACC
CAAAGAACATCCCCCTGAAATCATTATACCAAGAATATCATTTACCTGATTTACATCAAAATGAAACAACCTCTGAAAGAATCTTCAAGAAGATCCTTGA
ATTCATGCTGAAGCCCTCTGTGAAAGCTTCAAGGATGGGGCAAAAGGCCATTGAAGTGGCTCTTGCATCAGGG
TTCTATAGCTTTGAATGTACACGGACACCCAGGATGGGGGTGACAAGCTGTTTAC
TTCTGGAAAGTGGCTGAAGGGCAATTAAAGAAGCTGAGATTCCGTGACCCGGGGTGTGACAAAGCTGTTTAC
TCTAACCCCTGGCAATTCAACACGGCTTACTTCATGACACTTGGAAAGACTTGGCAAAAGCTCCAAAGCAATTGATGT
CTAA

>PAK5_h
62/76 C G G G A G

SUBSTITUTE SHEET (RULE 26)

Fig. 1E

63/76

64/76

GCAGCAAAGTGGAGAA, GATGGAGCAAGACCATGCCGTGGCGGGATCCGCCGGTGGAGGAGC
GGGACTACACCAGTTCCAAGAGCAGCTCAAACACTGATGAAGAAAGAGGTGAAGAACGAGGTGAGAAGCTCCCCGACAG
CAGCGGAAGGAAGCATGAAGCAGAAAGATGGAGGACACCGAGAACGGACACGGAGCTGGCCATTGACCGGGACTTTGTAGCCAA
GCGAGAGGAGCACCTGGAGCTGGAGCTGGCCATGAAGAGGAGCTGGAGGACACGGAGAACGGGAGATCTGTGACAAGGAGC
GCCTCATGAAGAACGGAGCTCCCTTGAGACCCGGAGGAGCTGGGAGATGGAAAGAGCACCCAGCTGAGGAGG
CACCAAGCTGGTGAAGCAGCAAGCTCAAAGACCAAGTACCTCCAGGGCACGGAGCTGTCGCAAGCATGAGAAGGAGCG
GGAGCAGATGCAAGGGCTACAAACAGGGCATGATGAGCAGCTGAAAGGTGCGGAGCAACAGGAAAAGGGGGCTGCCCA
AGATCCAGAGGAGTGAAGGCAAGACGGCATGGCCATGCCATGTACAAGAAAGGGCTCCACATCAACGGGGCAAGGGAGCT
GAGCAGGCTGAAGAGATCAAGCAGTTCTCCAGCAGGAGGAAGAGGGAGAAGTCTGGAGGAGCTGCAAGAACAGCAGAA
ACACGAGAACCAAGATGCCGGACATGCTGGCAGTGGCAGAGCAACATGAGGGAGCTGCAAGAACCTGAAGGAATGGGG
GCCACCTCTGGTAGAGCACGAAACCCAGAAACTGAAGGGCTGGAGAGGATCTGAACCCAGAAGGAAGGGAGCTTCAAGCT
GACAAGCTTGGGGCAAGAAGGCTCTGGAGAGGAGATCTGAACCCAGAAGGAAGGGAGCTGGAGGAGATGGGGATGCTT
65/76 CC

Fig. 11H

STLK6_h	- - - - - M S L D C F C T S R Q V E S L R P E K Q S E T S I H Q Y L V D E P T L S W S R	41
STLK5_h	- - - - -	0
SPAK_h	W A E P S G S P V H V Q P Q Q A A P V T A A A A A P A A A T A A P A P A P A P A P A	50
STLK6_h	P S T R A S E - V L C S T N V S H Y E L Q V E I G R G F D N L T S V H L A R H T P T G T L V T I K I	90
STLK5_h	- - - - -	
SPAK_h	P A A Q A V G W P I C R D A - - - Y E L Q E V I G S G A T A V - - - V Q A A L C K P R O E R V A I K R	95
STLK7_h	- - - - - N R D D Y E L Q E V I G S G A T A V - - - V Q A A Y C A P K K E K V A I K R	35
STLK6_h	T N L E N C N E E R L K A L Q K A V I I L S H F F R H P N I I I Y W T V F T V G S W L W V I S P F M A	140
STLK5_h	T N L E A C S N E M V T S - C R A S C M F Q T L N H P N I V P Y R A T L I A D N E L W V V T S F M A	58
SPAK_h	T N L E K C Q T S M D E L L K E I Q A M S Q C - S H P N V V T Y Y T S F V V K D E L W L V M K L L S	144
STLK7_h	T N L E K C Q T S M D E L L K E I Q A M S Q C - H H P N I V S Y Y T S F V V K D E L W L V M K L L S	84
STLK6_h	Y G S A S Q Q L R T Y F P E G - - - - - M S E T T L I R N I L F G A V R G L N Y L H Q N G C I H R S	184
STLK5_h	Y G S A K D L T I C T H F M D G - - - - - M N E L A I A Y I L Q G V L K A I D Y I H H M G Y V H R S	102
SPAK_h	G G S M L D I I K Y I V N R G E H K N G V L E E A I A T I L K E V L E G L D Y L H R N G Q I H R D	194
STLK7_h	G G S V L D I I K H I V A K G E H K S	103
STLK6_h	I K A S H I L I S G D G L V T L S G L S H L H S L V K H G Q - - R H R A V Y D F P Q F S I S V Q P W	232
STLK5_h	V K A S H I L I S V D G K V Y L S G L R S N L S M I S H G Q - - R Q R V V H D F P K Y S V K V L P W	150
SPAK_h	L K A G M I L L G E D G S V Q I A D F G V S A F L A T G G D V T R N K V R K T E - - - V G T P C W	240

S P E L L	R Q D L H G Y N V K S D I Y S V G I T A C E L A S G Q V P F Q D M H R T Q M L L Q K L K	282
S P E V L	R Q D L H G Y N V K S D I Y S V G I T A C E L A S G Q V P F Q D M H R T Q M L L Q K L K	200
M A P E V M E O - V R G Y D F K A D M W S F G I T A I E L A T G A A P Y H K Y P P M K V L L M L T L Q	289	

STLK6_h	P S - S[K T F S P A F F S [L V Q L C L Q D P E K R P S A S S L S H V F F K Q M K E E S Q D S I L 381
STLK5_h	P Y - H R T F S P H F H F V E Q C L Q N P D A R P S A S T L L N H S F F K Q I K R R A S E A [P 298
SPAK_h	K E M M K Y G K S F R K L L S L C L O K D P S K R P T A A E L L K C F F Q K A K N R E Y - - L I 349

STLK6_h	SLLPPAYNKPSISLPPVYL ELLRPV - - TPI	SIISLPPVYL GSDQHSGIFGLVTN RVPGS	KDSYWEF DDWEE VDDWEF	- - ECDFPDEK - - SDDDEMDEK	418 338 394
STLK5_h	EKLRT SPAK_h				

SPAK **SEE** GKAAFSQEKSRRVKEENPEIAVSASTIPEQIQSSLSVHDSQGPPNAME 444

DYREASSCAVNLLVRNRSRKELNDIRFEFTPGRDTADGVSQELFSAGLV 494

SPAK h DGHGVVIAANLOKIVDDPKALKTLLTFKLASSGCCDGSEIPDEVKLIGFAQL 544

547

Fig. 12B

PAK1_h U51120 MSNNGLIDIQDKPPAPPVRNTSTMIGAGSKDAGTLNHGSKPLPPNPEEKKKDRFYRSILP 60
 PAK4_h ----- 0
 PAK5_h ----- 0

PAK1_h U51120 GDKTNKKKEKEPEISPSDFEHTHVGFDAVTT-GEF|G|P|E|Q|W|A|R|L|Q|T|S|N|I|T|K|S|----- 115
 PAK4_h ---MF|R|KKKKRPEISAPQFQHRVHTSFDEEGKFYGLPPQWQNTLD-T|T|R|RPKPVVDP 56
 PAK5_h ---MF-GKRKKRVEISAPSNEHRVHTGFDQHEOKFTGLPRQWQSTEEESARRPKPLVDP 56

PAK1_h U51120 SRITTRYQCPWIKTVVRGSAWPDGYESGLIND|QK|SVISSNTILRG|SPTSRRAQSLSGL 115
 PAK4_h ACITSQOPCAPIKTIYRGSKGAKDGAFLIDEENMSVTRSNISLRRDSPPPPARAR----- 116
 PAK5_h ----- 112

PAK1_h U51120 LGDEHWATDPDMYLQSPQSERTDPHGLYLSCNGGTPAGHKQMPWEPQSPRVLPNGLAAK 115
 PAK4_h ----- 176
 PAK5_h ----- QENGMPPEEPATTARGGPCK 131

PAK1_h U51120 AQS|L|GPAAEFQGASQRC|L|QLGACLQSSPFGASPPPTGTNRHGMKAAKHGSEEARPQSCLYGS 127
 PAK4_h AGSRGR----- 236
 PAK5_h ----- QAGHSEA GGGSREGSGGPQESSRDKRP-----LSGR 183

PAK1_h U51120 EFKKNPOAVYDV 127
 PAK4_h ----- 174
 PAK5_h ----- 296

PAK1_h U51120 EFY-NSKKTSNSQKYMSFTDKS----- 174
 PAK4_h ATGRPGEGSPKTRESTSLKRLFRSMFLSTAAATAPSSSKPGPPQSKPNSSERPPQK 296
 PAK5_h DVGTPOPAGLASSAKLAG----- 232

PAK1_h U51120 E D D D D D D A T P P P V I A P R P E H K S V Y R S - - - V I E P L P T P R D N A T S P T E N N I T 230
 PAK4_h D N P P S L V A K A Q S P D Q P V G T F S P L T S D T S S P Q K S E R T A P A T G Q L D G R S S P A G S P R T W H 356
 PAK5_h P O L A P P A - C T P A A P A V B G P P G P R S P O R K E P O R Y S H E Q F R A A L Q 268

PAK1_h U51120 P - - - - - P D A C T L N T E K Q K K P K T S D E E L E K L R S E V S G D P K K K Y T R F E K I 276
 PAK4_h A Q I S T S N L L P Q D P T V A K G A - - - L A G E D T V W T H E Q F K A A L R M V V D Q D P R I L L D S Y Y K I 413
 PAK5_h P O L A P P A - C T P A A P A V B G P P G P R S P O R K E P O R Y S H E Q F R A A L Q 327

PAK1_h U51120 G Q G A S G J I V Y T A M D V A I G Q E V A I K Q M N L Q Q Q P K K E L I T N E I T M R E N K N P N E V N Y L D S Y L V 336
 PAK4_h G E G S T G I V C I A R E X H S G R Q V A V K M D I R K Q Q R E L L F N E V I M R D Y Q H E N V V E M Y K S Y L V 473
 PAK5_h G E G S T G I V C I A T V R S S G K L L V A V K K M D L R K Q Q R E L L F N E V I M R D Y Q H E N V V E M Y N S Y L V 387

PAK1_h U51120 G D E L W V V M E Y L A G G S T D D V T E T C M D E G Q I A A V C R E C I Q A L E S L H S N O V I H R D I K S D N I L 396
 PAK4_h G E E L W V Q M E F L Q M G A L T D I V S Q V R N E E Q I A T V C E A V L Q A L A Y L H A Q G V I H R D I K S D S I L 533
 PAK5_h G D E L W V V M E F L E G G A L T D I V T H T R M N E E Q I A A V C L A V L Q A L S V L H A Q G V I H R D I K S D S I L 447

PAK1_h U51120 L G M D G S V K I L D F G F C A Q I T P E Q S K R S T M V G T P Y W M A P E V V R K A Y G P K V D I W S L G I M M A I E 456
 PAK4_h L T L D G R V K L S D F G F C A Q I S K D V P K R K S I L V G T P Y W M A P E V I S R S Y A T E V D I W S L G I M V I E 593
 PAK5_h L T H D G R V K L S D F G F C A Q M S K E V P R K S I L V G T P Y W M A P E D I S R L P Y G P E V D I W S L G I M V I E 507

PAK1_h U51120 M I E G E P P Y L N E N P L R A L Y I A T N G T P E L Q N D E K E S A T E R D F L N R C L E M D V E K R G S A K E L L 516
 PAK4_h M V D G E P P Y F S D S P V Q A M K R D S P P K N K N S H K V S P V L R D F L E R V R D P Q E R A T A Q E L L 653
 PAK5_h M V D G E P P Y F N E P P L K A M K I R D N P P R L K N H K V S P S L K G F L E R L V R D P A Q R A T A A E L L 567

545 681
 591

PAK1_h U51120 QHQQFLKIAKPKSSSLIPLIAAATK
 DHPFLQTGPPPECLVPLIQYRKQTS
 KHPFLAKAGPPASIVPLRQNRTT R

PAK4_h

PAK5_h

RESIDUES THAT MATCH THE CONSENSUS NAMED CONSENSUS #1 EXACTLY.

BOX RESIDUES THAT MATCH THE CONSENSUS EXACTLY.

Fig. 13C

ZC4_h.pro	MAGPGGWRDREVTDLGLPDPPTGIFS	LDTI	GLGTYGRIVLGEHEKTGAF	50	
ZC1_h.pro	MANDSPAksLVDid	SSLRDPAGIFELVEVVGNGTY	GQVYKGRHVKTGQL	50	
ZC4_h.pro	TAVKV	MNARKDEEDERTELNLIRKYS	SFKNNIVSEYGAFFKLSPGQRH	100	
ZC1_h.pro	AAIKV	MDVTEDEEETIKLEINM	KKYSHHRNATYYGAFIKKSPPGHDDQ	100	
ZC4_h.pro	LWMVME	LCAAGSYTDYVRMTSNOSLKE	DWIAYICREILQGLAHLHAHRY	150	
ZC1_h.pro	LWLMF	CAGSITDVKNTKGNT	EDWIAISREILRGLAHLHHRV	150	
ZC4_h.pro	HRDIK	GQNYLTTHNAEVKLYDFGVSAQ	VSRITNGRRNSFIGTPYWMAPENV	200	
ZC1_h.pro	HRDIK	GONVLTENAEVKLYDFGVSAQ	LDRTVGRNTFIGTPYWMAPENV	200	
ZC4_h.pro	DCDEDPPRSYD	YRSPDVWSYGITATEMAEGAPP	LCNLQPLEALEFYJIRESA	250	
ZC1_h.pro	ACDENPDATYD	YRSDLINSCGITA	TEMAEGAPP	CDMHPMRALFLIPRNPP	250
ZC4_h.pro	PTVKSSGWSRK	FHMFMKCTIKNELLFRPTSANML	QHPPFYRDIKNERHVY	300	
ZC1_h.pro	PRLKSKKWSKK	FFSFTEGCVKNYMQRPSTEQ	QLPKHPFIRDQPNERQVR	I	300
ZC4_h.pro	SLTRH	LTGIIKKR	-OKKEQAREKKKS	324	
ZC1_h.pro	QLKD	IDRTRKKGKDET	EEYESGSSEEEEEVPEQE	350	
ZC4_h.pro	KYSTLROALA	KRLSPKRRAKS	SSWRPEKLEISDLEARRO	RQRREDIFN	374
ZC1_h.pro	TLRRDFLRL	QQENKERSEALRRQ	QLLQEQQREQEEYKRQLLAER	RQKRIE	400
ZC4_h.pro	QHEEE	ERQVDDKDKED	ESSDNDEVH	HIOAEVQIEPLKPYISNDKKIEVDE	424
ZC1_h.pro	QKEQ	QRRRLEQQRREARRQ	QEREQRRREARRQ	EKRRLEELERRKEEEE	450

ZC4_h.pro RSPSWPNNODDHAHHVKEFSSSVPORSLLIEDAOKPIDIRQRSQNROQNWLA 474
 ZC1_h.pro RRRAEEKRRVEREQYIRRQLEEEQRHLEVQQQLLQEQAHLLECRWRE 500

ZC4_h.pro SGDSKHKKILAGKTTOSVCLTYISEVZKKEEFOEGMNOKC0GA0VGLGPEGH 524
 ZC1_h.pro MEEHRQAERLQRQLQQEQAYLLSILQHDHRRPHPQHSQQQPPQQERSKPS 550

ZC4_h.pro C1WQLGESSSEESPVITGRRSQSSSPPISTTIDQKLLVDIHWPDGFKYGGKIS 574
 ZC1_h.pro FHAPEPKAHYEPAADRAREVEDRFRKTNHSSPEAQSKQTGRVLEPPVPSRS 600

ZC4_h.pro PPyYLINENWGYNALSEIIFRNDSLTPAPVIOOPPEEDGDDYWELEYDASADID 624
 ZC1_h.pro EFSNNGNSESVHPALQRPAEPQVPUVRTTSRSPVLSSRRLQGSQQQNSQ 650

ZC4_h.pro GDDDESNDTFEEDIYDHANGNDLDLNDNQVDOANDVCKDHDNNNKTVDDVN 674
 ZC1_h.pro AGQRNSTSIEPRLLWERVEKLVPVRPGSGSSSGSSNQGSHPGSQSGS 700

ZC4_h.pro NNYYEAPSICPRASYGRDGSSCK0DGYDLSRGKEEAYRGYGSHTANRSHGGGS 724
 ZC1_h.pro GERFRVRSISKSEGSPSQRLENNAVKKPEDKKEVFRPLKPADLTALAKER 750

ZC4_h.pro AASEDNAALGDQEEHANIGSERRGSSCGGGKGIVRTSEESSGALLINGEE 774
 ZC1_h.pro AVEDVRRPPHKVTDYSSSSSEESGTTDEEDDVEQEADESTSGPEDTRAAS 800

ZC4_h.pro NCSEIDGPICIKRPAASQDFEYLOEEPGCNEASNAIDSAAAPDHESDN 824
 ZC1_h.pro SLNLSNGETESVKTWIVHDDVESEPARNTPSKEGTLIVRTQSASSTLQKH 850

ZC4_h.pro KDISS9ST05DEFSAHHSSSPSKGGSGMSADANEASALLYAGIVEVPPESPKQ 874
 ZC1_h.pro KSSSSFTPFIDPRLLQISPSGTTVTSVVGESCDGMRPE--AIRQDPTRK 898

Fig. 14B

Fig. 14C

BOX RESIDUES THAT MATCH ZC4_h.pro EXACTLY.

74/76

SUBSTITUTE SHEET (RULE 26)

Db	361 S - STPL PPSQP QEPVNGPCSQPSQDLSKNDNDLKV 361 SPSTPL APSQS QDSVNEPCSQPSQDRSLQTTSPVPLRKS Qy	VPLRKS RPLSMSDA 419 VPLRKS RPLSMSDA 420
----	--	--

540 SITTSKIISEDEKKDEEMRFLRRQELRELRLQQKEEHRNQTQLSSKHELQLEQMHHKRFEQ 599
541 SITTSKIISEDEKKDEEMRFLRRQELRELRLQQKEEHRNOTOLSNKHELQLEQMHHKRFEQ 600

* * * * * 660 KKEVKSEVEKILPRQQRKESMKQKMEEHSSQKKQRLDRDFVAKQKEDLELAMRKLTENRRE 719
* * * * * 661 KKEVKNEVEKILPROQRKESMKQKMEEHHTOKKKQLLDRDFVAKQKEDLELAMKRLTTDNRRRE 720

15B

76/76
SUBSTITUTE SHEET (RULE 26)

Fig. 15C

SEQUENCE LISTING

<110> SUGEN, INC.

<120> STE20-RELATED PROTEIN KINASES

<130> 240/300-PCT

<140> TO BE ASSIGNED

<141> Herewith

<150> US 60/081,784

<151> 1998-04-14

<160> 147

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 3268

<212> DNA

<213> Mammalian (Human) STLK2

<400> 1

taacagccca	cctccttagcc	ccgggctacg	cgccgccagc	ccagtaaccc	cactttgtg	60
tgtcctccca	ggccccgatc	gaaaagcctg	ggaggggccgc	cgaactaccc	ccggaggggag	120
gagccagtcc	gaacccaagg	cgccaccgc	gcagaagcgg	agcgaggcag	cattcgcc	180
catggcccac	tcgcccgtgg	ctgtccaagt	gcctggatg	cagaataaca	tagctgatcc	240
agaagaactg	ttcacaaaaat	tagagcgtat	tggaaaggc	tcatttgggg	aagtttcaa	300
aggaatttgat	aaccgtaccc	agcaagtctgt	tgctattaaa	atcatagacc	ttgaggaagc	360
cgaagatgaa	atagaagaca	ttcagcaaga	aataactgtc	ttgagtcaat	gtgacagagtc	420
atatgtaa	aaataactatg	ggtcatattt	aaagggtct	aaatttatgga	taataatgga	480
atacctggc	ggtggttca	cactggatct	tcttcgagct	gttccatttgc	atgagttcca	540
gattgctacc	atgctaaagg	aaattttaaa	aggcttgac	tatctgcatt	cagaaaagaa	600
aattcaccga	gacataaaaag	ctgccaatgt	cttgctctca	gaacaaggag	atgttaaact	660
tgctgatttt	ggagttgctg	gtcagctgac	agatacacag	ataaaaagaa	atacctttgt	720
gggaactcca	tttggatgg	ctcctgaagt	tattcaacag	tca	actcaaaaagc	780
tgacatttgg	tcattggaa	ttactgctat	tgaactagcc	aaggagagc	cacctaactc	840
cgatatgcat	ccaatgagag	ttctgtttct	tattccaaa	aacaatc	caactcttgt	900
tggagactt	actaagtctt	ttaaggagtt	tattgatct	tgc	ctgaaca	960
attcgtcct	acagcaaaaag	aacttctgaa	acacaattc	ttgtaaaaaa	attcaaagaa	1020
gacttcttat	ctgactgaac	tgatagatcg	tttaagaga	tggaggcag	aaggacacag	1080
tgtatgatgaa	tctgattccg	agggctctga	ttcggaatct	accagcaggg	aaaacaatac	1140
tcatcctgaa	tggagcttta	ccaccgtacg	aaagaaggct	gatccaaaga	aagtacagaa	1200
tggggcagag	caagatcttg	tgcaaaccct	gagttgttg	tctatgataa	tcacac	1260
atttgctgaa	cttaaacagc	aggacgagaa	taacgctacg	aggaatcagg	cgattgaaga	1320
actcgagaaa	agtattgctg	tggctgaagc	cgccgtcccc	ggcatcacag	ataaaaatgg	1380
gaagaaacta	attgaaaaat	ttcaaaagtg	ttcagcagac	gaatccccct	aagaaactta	1440

ttattggctt ctgttcata tggacccaga gagccccacc	aaacctacgt caagattaac	1500
aatgcttaac ccatgagctc catgtgcctt ttggatctt	gcaacactga agatttgaa	1560
gaagctatta aactatttt tgatggcggtt tatttttta	tattttgaaa ggattatttt	1620
gtaaggaaata acttttaata ctatagttc acctgtattc	tagtaaatgt tgagacaccg	1680
tttgctttt aagtatccct atttcttaag ttacgaggat	gaatacctt cacattttga	1740
tcttttagttg actctacagt catgaacat acaggtcttt	caaagtctt ctaaatattc	1800
agcttttgta aatttatcaag cttcaaaaag cttttttta	aaaaaaaaaa catgcatatt	1860
ctaaaaatga ctattggtgg ggagggttaa ataagtctata	ccttcttaaa acagaaaatt	1920
taagtaaatgt cttttaatgt aaacctgtaa aagtattgac	tcttctacca agttggatg	1980
atattccagg cagctcaatg attatcacat ttgagaccct	gtgttgaag catttacagg	2040
caatgtacag caacagaggta accttttgtt gtatagtatt	tacattctt ttaggtaga	2100
agaggcaatt ttacccttat ttcacatggt tagaaattta	aagcaagatc atttacccaa	2160
ggatagggtgt ttggtaatgt tgaaggagtt agtctggctt	catgttttac atcttcaact	2220
aaaatcccat actatctgct tggatttggg gagccaaaaa	ataaaagctga ttgtcatgtg	2280
attaaatatac tcatcaacag gtatgaatat aacttaatc	agcatattt tgccatggta	2340
ataaaattgtc ctataaacta tttatattt tttgttcttc	ataattatca ctaataagca	2400
tcagttgtt gttttaaaaa ggatatttaa gtgagcattt	tctagttcat atgaaaataa	2460
ccatagtaca ggatgatttgc tgtccacaca aaggtaaat	tagattgcac agttaatttt	2520
cacttatatt tatggtaact ttatgtgggt gatgccttt	tcttttaagc ccagtagata	2580
tattatgcct gcctaagttc tgaactgggg ctgtatttca	gtagttgtag aattattgtat	2640
atttagttt gatagctaattgtt gtttgcatttgc acagtttgtt	ttttgcacaa	2700
aagtcatat aaaaatctg agtaattgtc aaatattaaa	agaaagatat tttcctgt	2760
aggaatacag tttttagtca aagtggccat tacatcctt	tttttaattt cataatacag	2820
atacttgaga aagttgtgtt ggtgtgtat gccaagaaaa	ttctttttat tggtgcttat	2880
attgtaacaa ttatttttaa tgcattgtat tttgaagtaa	cggtcagtt aaatttttca	2940
cctgctgtgt aactgaaaca caattacagt ttataatcat	ctgtagaagt ctggagataa	3000
ttttgcaact catgttatgg gttaaatgaa tattttttaa	aaagtaaaaag caacaatatt	3060
ataaaattgtat tatttgcacac aattgcattcc caaatacaaa	ttgtattgt	3120
tattcattat agctattcgt cctgtatct gtttcttagt	gaagcatact ccagtgttt	3180
aggggttttggaaaataata tttaaatttc acgtcaaaa	aaaaaaaaaa aaaaaaaaaa	3240
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa		3268

<210> 2

<211> 3030

<212> DNA

<213> Mammalian (Human) STLK3

<400> 2

gacagcagcg ccggccccgg cagctcccgc ggccccggcc	ccggccccgg ccccgcccc	60
ggcggcacag gctgtcggtt ggcccatctg cagggacgc	tacgagctgc aggaggttat	120
cggcagtggc gctactgctg tggttcaggc agccctatgc	aaacccaggc aagaacgtgt	180
agcaataaaaa cggatcaact tggaaaaatg ccagaccagt	atggatgaac tattaaaaga	240
aattcaagcc atgagtcagt gcagccatcc caacgtatgt	acctattaca cctctttgt	300
ggtcaaaatgtt gaaactttggc tggtcatgaa attactaagt	ggaggttcaa tggatgtat	360
cataaaatac attgtcaacc gaggagaaca caagaatgg	gttctggaaag aggcaataat	420
agcaacaatt cttaaagagg ttttggaaagg cttagactat	ctacacagaa acggtcagat	480
tcacagggat ttgaaagctg gtaatattct tctgggtgag	gatggttcag tacaatagc	540
agattttggg gtaagtgcgt tcctagcaac aggggggtgat	gttacccgaa ataaagtaag	600
aaaaacatttcc gttggcaccc catgttggat ggctcctgaa	gtcatggaa aggtgagagg	660
ctatgacttc aaggctgaca tggatgttt tggataact	gccattgaat tagcaacagg	720
agcagcgcct tatcacaat atccctccat gaaagtgtt	atgttactt tgcaaaatga	780
tccaccact ttggaaacag gggtagagga taaagaaatg	atgaaaaagt acggcaagtc	840
cttttagaaaa ttactttcac tggatcttca gaaagatctt	tccaaaaggc ccacagcagc	900

agaacttta	aaatgcaa	at tcttcagaa	agccaagaac	agagagtacc	tgattgagaa	960
gctgcttaca	agaacaccag	acatagccca	aagagccaaa	aaggtaagaa	gagttcctgg	1020
gtcaagtgg	cacccata	aaaccgaaga	cggggactgg	gagtggagt	acgacgagat	1080
ggatgagaag	agcgaagaag	ggaaacgcgc	ttttctcg	gaaaagtca	gaagagtaaa	1140
agaagaaaat	ccagagattt	cagttagtgc	cagcaccatc	cccaccaaa	tacagtcct	1200
ctctgtgcac	gactctcagg	gcccacccaa	tgctaatgaa	gactacagag	aagttcttc	1260
ttgtgccgt	aacctcg	ttt	tgagattaag	aaactccaga	aagaactta	1320
atttgagttt	actccaggaa	gagatacagc	agatggtga	tctcaggagc	tcttctc	1380
tggcttgg	gatggtcacg	atgtagttat	agtggctgct	aatttacaga	agattgtaga	1440
tgatcccaa	gctttaaaaa	cattgacatt	taagttggct	tctggctgt	atgggtcgga	1500
gattcctgat	gaagtgaagc	tgattgggtt	tgctcagg	agtgtcagct	gatgtatgtc	1560
ccttgatgtc	accctgatct	gtcatgcccc	accgcaccc	ctactccctt	caacccccc	1620
tcttctgcc	catttctcc	cacccctca	ctcccttcc	ctagcaaaaat	cagaagattt	1680
tgaagaggcc	ggcttcaaca	aatggata	aaaaaataat	ttttaaaac	ttacaacact	1740
ccgagttctg	cttattctc	tagcaatcca	cagtacaaga	acaagcaat	gccacagctg	1800
cacgactgtt	gctcattttt	ccaaaagcta	ttaatattc	ttagcaatca	atttggat	1860
cccttaagt	aaaagaatct	gaaatacact	caggtggct	tatttatttg	caacaaaagg	1920
aattttctat	ccagaagcct	atttctc	tcttattgt	tatttctgtt	ataatactt	1980
aattgtacat	ctgacaatac	tgcctt	atgttgtatt	tagaaattaa	tataacttata	2040
aaattaagat	ttattagcca	aacttgaatt	ctagttttaa	aactgactgt	gaattttatt	2100
tttcatatat	ttatgcatta	cacaccc	ctataagaaa	aaaagggtt	tgatttat	2160
cttcttg	cttgcag	ttaatctcg	tattttaaca	aaaagttt	ggtctatctt	2220
gttaacttct	aattttgaaa	tgactgaatt	aggaatttgg	atgttattc	tttttgtctg	2280
tttgcctaaa	aaccaattt	caatctgact	gtcttgg	agagggaggt	gccttgcaaa	2340
cttccacatt	aagaatgtc	ctgagctgc	tttactctgg	aatagtctca	gatctaaaat	2400
ttccctata	taaggtggca	tatgttaagt	tttgc	tggaccgtt	agaatgtat	2460
gtaaaatgtt	gccattctgt	tagattgct	actatatacc	catctctgat	ttggctctcc	2520
ttaagtgata	ggatttggta	ttctaaaggt	gataaaactt	aaaatatcag	aatctgagtt	2580
ttacttgaaa	tttgcagaa	tacccaggt	gagtggaaaat	tggagggtt	ttgtcaatg	2640
actaaaaggt	aaaacgctgt	taaggtcaa	gaatcaatac	tttcaaccc	agtagccctc	2700
tgcttgact	tatattatgg	aactagtaaa	ccttaggatt	ttggaaaattt	gagtctaa	2760
tttcaaggag	gtggctccc	aggatgtac	cattgctt	tcctagctaa	ccctagat	2820
ggcagctt	taatgtactt	caaaaagcaa	atataatatta	ctaaggaaa	aaagttat	2880
ataattgcct	tgtcataatt	gttaaggtgt	tctagagcca	tttgcataca	atthaatgt	2940
atttcatcc	atttattgt	ttacacaacg	attactcgaa	gatgactgca	aaggtaaaag	3000
aaaaataaaa	gtgtattgca	caatggaaaa				3030

<210> 3
 <211> 3857
 <212> DNA
 <213> Mammalian (Human) STLK4

<400> 3

caaaaagtgg	gtcctagat	antctaccat	tgctacgata	ctccgagaag	tactggagg	60
gctggat	at	ctgcataaaa	ntggacagat	ccacagagat	gtgaaagct	120
tnttggagaa	gatggct	tagattt	agacttgg	tttagtgctt	tttagcaac	180
tgggtgtat	attaccc	gaa	aaagacctt	tttggcaccc	cttggat	240
ggcacctgaa	gttatggaa	agg	ttatgat	aaagctgata	tttggat	300
tggattaca	gcaattgaat	tgg	ctacagg	ggcggct	tatcataat	360
gaagg	ttt	atgctgac	tgc	caac	agac	420
taaagaaat	ctgaaaaat	atggaaaat	at	gat	tttca	480
aaaagat	caaaaaagac	caac	agc	aga	actt	540
agcaaa	aaagaattt	ttcaagaaa	aa	acatt	ggc	600

aagagcaaaa	aagggtcgga	gagtaccagg	ttccagtggg	cgtcttcata	agacagagga	660
tggaggctgg	gagtggagtg	atgatgaatt	tgatgaagaa	agtggaggaag	ggaaagcagc	720
aatttaccaa	ctcaggtctc	cccgagtgaa	agaatcaata	tcaaattctg	agctcttcc	780
aacaactgat	cctgtggta	cttgcgtcca	agttccagaa	cagatctctg	ctcatctacc	840
tcagccagct	ggcagattt	ctacacagcc	aactcaagtc	tctctccac	ccaccgcaga	900
gccagcaaaa	acagctcagg	ctttgtctc	aggatcaggt	tcacaagaaa	ccaagatccc	960
aatcagtcta	gtactaagat	taaggaattc	caaaaaagaa	ctaaatgata	ttcgattga	1020
atattactct	gggagagata	cagcagaggg	tgtctctcg	gaactcattt	ctgctggcct	1080
ggtcgacgga	agggatttag	taatagtggc	agctaattt	cagaaaattt	tggagaacc	1140
tcagtcaaatt	cgtctgtca	ctttcaact	ggcatctgtt	gtcgaaggct	cagatattcc	1200
tgatgttgtt	aaactgtatag	gatttgcctc	gctcagcata	agctaaacca	caaccctgga	1260
agaggcggcc	taaggagatt	ccacacatgc	gtatctctgt	tgcttctatt	ggcctaaacc	1320
caactgtcc	aaagaaccca	gcaacaaacc	tcccggctag	gagctttaga	agtctttatg	1380
ttcttcctgc	catcattcc	cctttccca	cagggaaaga	aaagttggat	caactgtggc	1440
cagcatcccc	agagttccgt	tagtaaactt	acttcataatg	tcccctgtct	tcctccatct	1500
gagaagtggc	ccatgtgctt	caaggcccag	gagggagatc	tgtcagctca	ttcttcgcctt	1560
actccaatga	tggcccagg	ggaaaagtag	cagctgtatc	ggccttcctc	atcctgcctg	1620
ttccccccaca	cctgcccagg	tatggacatc	ttgggatatc	tcttaccac	tgaagtagaa	1680
ttgatttttc	agctggagcc	cagagaattt	aatttaatgt	tttttctttt	tacctgatgt	1740
gaattcttagc	aacctttgtt	aggaaaaagc	acagcctcg	atggaggcag	cctaaactgt	1800
gttcttggtt	tgttcatgg	gtttctaagc	gttttgcgtga	agtcgtctc	aggcaccccc	1860
ttcttcattt	ctctctccag	aaagggttgc	tagccttaac	ttcagctgg	gcaaaacatc	1920
tgactgttagc	cgaacttcag	ccatcagatc	cttcaaagt	gaacttttgg	ttgtttttac	1980
agacaacatc	gagtaatggc	ttgtaaaatgt	gaattttgc	agaggtgg	tttgaacagg	2040
aaaatcataa	ttcatatcat	tggagaagta	tttattttca	aatatcaa	tgaagaaaaa	2100
ctcaatcc	ccatgaaaat	cagttcgct	ggccctccaa	tcgtgaggaa	atgggtatgc	2160
aaggctgaga	tttctacagc	aataaaggag	acacacactg	ggccagagag	gcctgcctc	2220
tgcctgctt	cctgcactga	ccctttggag	ggggctctg	tgtctgaag	ctaaactcaag	2280
atggaaagt	aaaccacatg	tgccgtgacc	tttaggttt	atgagtagac	agtgttcatt	2340
tgattttcta	cagaataat	ataaattatt	ctttaggtt	aaaaaaagagc	actcataatg	2400
caatatgtga	ataatcagtg	aggttgattt	ttctttttc	ctaccgtt	atagtcttt	2460
tctaactgct	agtaacccta	ccgagtttta	tatatgagtg	ggatactcaa	tctggcctt	2520
aaaagataca	caaagatggg	ctgtgggtcc	ctggaaaggg	ggagagttgc	cctttacaga	2580
atcactcgag	ccctttccag	cactgttgtt	ctgatgaaca	aggttgg	accttatttt	2640
cttttggaa	atatactgaaa	accttccca	caaataactt	gtcacac	ttgtttcatt	2700
ctgagtttt	agtttttagt	atggctttc	ttcacctgt	ctagg	gcata	2760
ggaaagagat	ggatgttggg	gaggaagaga	ggagatgg	ttcagttgg	agtaggagg	2820
agagtaggt	agatgtatcg	acaccggagt	tcaacgtcc	agcagtctt	gtaaaaggag	2880
ggagcctgct	gagccaggag	ggagaaaaa	agattgacca	gctgtctaga	aaaatactta	2940
gtttttctt	ttttttttt	gtggaggggg	gacggagagg	aacaaggatg	gggaggtagg	3000
aatgaggtat	agaaaagaga	tagcatctc	tttggcaca	gactagtgg	ttaccgctt	3060
ccttagagtt	ttgtttttt	ttttcaaa	ccatcaaaat	ctacttattt	atgaatccaa	3120
gggggtggcag	catca	ttctagcatt	cttgcgtt	atggcttgg	gcctagctgg	3180
gagtgacgac	cagccatcc	cctgttact	ttctctagcc	catcattacc	tgtgaactgc	3240
agtggggcag	tcatggcaaa	tagatttgg	ctggggttt	tccttctttt	cagttcattt	3300
tttgcctgc	taggaattt	aagacagaca	ccatgtccca	ggacagtgtt	acttcttctg	3360
catgatgtgt	ggtagactcc	ctttgtggc	ttgtgcgt	atactgagaa	aatacatgaa	3420
cagaaactgc	ccaggtggaa	cagcacgtaa	cctagtgtgt	gactgtactc	ctttcttagga	3480
atgctgattt	agagtgcacc	tcttgacta	ggtcccagga	tcccctgtc	cctggagtag	3540
ggactaacta	tagcaca	taatatgtgc	caatgctt	tgtgaaatgt	ttggctttt	3600
taaacgacta	aaggattt	tgggtttt	cttaagttt	gaaccaa	atcagagccag	3660
ctgataat	ttaataatct	ggaggagaga	ataatgtat	accaataatg	ggagattcc	3720
ccttatgtat	tatgtcttagt	tatgaaat	gtaaaatatt	caactttt	ctccttttt	3780
tggactttgt	atttactgc	atgtttctt	cattttat	caataaagag	taaattgtca	3840
aaaaaaaaaa	aaaaaaaaaa					3857

<210> 4
 <211> 1584
 <212> DNA
 <213> Mammalian (Human) STLK5
 <400> 4

ctcatctgta	cacacttcat	ggatggcatg	aatgagctgg	cgattgctta	catcctgcag	60
ggggtgctga	aggccctcga	ctacatccac	cacatggat	atgtacacag	gagtgtcaaa	120
gccagccaca	tcctgatctc	tgtggatggg	aaggcttacc	tgtctggtt	gcgcagcaac	180
ctcagcatga	taagccatgg	gcagcggcag	cgagtggtcc	acgatttcc	caagtacagt	240
gtcaagggtc	tgccgtggct	cagccccgag	gtcctccagc	agaatctcca	gggttatgtat	300
gccaagtctg	acatctacag	tgtgggaatc	acagcctgtg	aactggccaa	cggccatgtc	360
ccctttaagg	atatgcctgc	cacccagatg	ctgcttagaga	aactgaacgg	cacagtgc	420
tgcctgttgg	ataccagcac	catccccgt	gaggagctg	ccatgagccc	ttcgcgtca	480
gtggccaact	ctggcctgag	tgacagcctg	accaccagca	ccccccggcc	ctccaaacgg	540
gactcgcctt	cccaccccta	ccaccgaacc	ttctcccccc	acttccacca	ctttgtggag	600
cagtgccttc	agcgcaccc	ggatgccagg	cccagtgc	gcaccctcct	gaaccactct	660
ttcttcaagc	agatcaagcg	acgtgcctca	gaggcttgc	ccgaatttgc	tgcgcgtc	720
accccccata	ccaaatttga	ggcagccag	tctcaggacc	acagtggaa	ctttggctg	780
gtaacaaacc	tgaaagagct	ggaggtggac	gattgggagt	tctgagcctc	tgcaaactgt	840
gcgcattctc	cagccaggg	tgcagaggcc	acccagaggg	ccttcctgag	ggccggccac	900
attcccggcc	tcctggggcag	attgggtaga	aaggacattc	ttcaggaaa	gttgactgct	960
gactgattgg	aaaaaaaat	cctggagaga	tacttcactg	ctccaggct	tttgagacac	1020
aagggaatct	caacaaccag	ggatcaggag	ggtccaaagc	cgacattccc	agtcctgtga	1080
gctcagggt	cctcccccgc	agaagagaga	tgctgctctg	gccctggag	ctgaattcca	1140
agcccgagg	ttggctcctt	aaacccgagg	accggcacct	cttcccagt	cttgcgacca	1200
gcctcatct	attaacttt	gctctcagat	gcctcagat	ctataaggta	gtgaaaggc	1260
aagtagtaag	ctgctgcct	ccctccctc	agacctctcc	ctcataattc	cagagaaggg	1320
catttctgtc	tttttaagca	cagactaagg	ctggaacagt	ccatccttat	ccctcttctg	1380
gcttggggcc	tgacacctaa	gtcttccca	cggttatgt	gtgtgcctca	ttccttccc	1440
accaagaatc	cacccatcg	cctccctgca	gctgccctgg	tgcttctcc	aaggccatc	1500
agtgtcttgc	ctagcttgag	ggcttaagtc	ctttagctgt	gttagttcg	ttgtcagaac	1560
aaattaaaaat	tttcagagac	gctg				1584

<210> 5
 <211> 416
 <212> PRT
 <213> Mammalian (Human) STLK2

<400> 5

Met	Ala	His	Ser	Pro	Val	Ala	Val	Gln	Val	Pro	Gly	Met	Gln	Asn	Asn
1					5				10				15		

Ile	Ala	Asp	Pro	Glu	Glu	Leu	Phe	Thr	Lys	Leu	Glu	Arg	Ile	Gly	Lys
				20				25				30			

Gly	Ser	Phe	Gly	Glu	Va	Phe	Lys	Gly	Ile	Asp	Asn	Arg	Thr	Gln	Gln
				35				40				45			

Val Val Ala Ile Lys Ile Ile Asp Leu Glu Glu Ala Glu Asp Glu Ile
50 55 60

Glu Asp Ile Gln Gln Glu Ile Thr Val Leu Ser Gln Cys Asp Ser Ser
65 70 75 80

Tyr Val Thr Lys Tyr Tyr Gly Ser Tyr Leu Lys Gly Ser Lys Leu Trp
85 90 95

Ile Ile Met Glu Tyr Leu Gly Gly Ser Ala Leu Asp Leu Leu Arg
100 105 110

Ala Gly Pro Phe Asp Glu Phe Gln Ile Ala Thr Met Leu Lys Glu Ile
115 120 125

Leu Lys Gly Leu Asp Tyr Leu His Ser Glu Lys Lys Ile His Arg Asp
130 135 140

Ile Lys Ala Ala Asn Val Leu Leu Ser Glu Gln Gly Asp Val Lys Leu
145 150 155 160

Ala Asp Phe Gly Val Ala Gly Gln Leu Thr Asp Thr Gln Ile Lys Arg
165 170 175

Asn Thr Phe Val Gly Thr Pro Phe Trp Met Ala Pro Glu Val Ile Gln
180 185 190

Gln Ser Ala Tyr Asp Ser Lys Ala Asp Ile Trp Ser Leu Gly Ile Thr
195 200 205

Ala Ile Glu Leu Ala Lys Gly Glu Pro Pro Asn Ser Asp Met His Pro
210 215 220

Met Arg Val Leu Phe Leu Ile Pro Lys Asn Asn Pro Pro Thr Leu Val
225 230 235 240

Gly Asp Phe Thr Lys Ser Phe Lys Glu Phe Ile Asp Ala Cys Leu Asn
245 250 255

Lys Asp Pro Ser Phe Arg Pro Thr Ala Lys Glu Leu Leu Lys His Lys
260 265 270

Phe Ile Val Lys Asn Ser Lys Lys Thr Ser Tyr Leu Thr Glu Leu Ile
275 280 285

Asp Arg Phe Lys Arg Trp Lys Ala Glu Gly His Ser Asp Asp Glu Ser
290 295 300

Asp Ser Glu Gly Ser Asp Ser Glu Ser Thr Ser Arg Glu Asn Asn Thr
305 310 315 320

His Pro Glu Trp Ser Phe Thr Thr Val Arg Lys Lys Pro Asp Pro Lys
325 330 335

Lys Val Gln Asn Gly Ala Glu Gln Asp Leu Val Gln Thr Leu Ser Cys
340 345 350

Leu Ser Met Ile Ile Thr Pro Ala Phe Ala Glu Leu Lys Gln Gln Asp
 355 360 365

Glu Asn Asn Ala Ser Arg Asn Gln Ala Ile Glu Glu Leu Glu Lys Ser
 370 375 380

Ile Ala Val Ala Glu Ala Ala Cys Pro Gly Ile Thr Asp Lys Met Val
 385 390 395 400

Lys Lys Leu Ile Glu Lys Phe Gln Lys Cys Ser Ala Asp Glu Ser Pro
 405 410 415

<210> 6
<211> 516
<212> PRT
<213> Mammalian (Human) STLK3

<400> 6

Thr Ala Ala Pro Ala Pro Ala Ala Pro Ala Ala Pro Ala Pro Ala Pro
 1 5 10 15

Ala Pro Ala Pro Ala Ala Gln Ala Val Gly Trp Pro Ile Cys Arg Asp
 20 25 30

Ala Tyr Glu Leu Gln Glu Val Ile Gly Ser Gly Ala Thr Ala Val Val
 35 40 45

Gln Ala Ala Leu Cys Lys Pro Arg Gln Glu Arg Val Ala Ile Lys Arg
 50 55 60

Ile Asn Leu Glu Lys Cys Gln Thr Ser Met Asp Glu Leu Leu Lys Glu
 65 70 75 80

Ile Gln Ala Met Ser Gln Cys Ser His Pro Asn Val Val Thr Tyr Tyr
 85 90 95

Thr Ser Phe Val Val Lys Asp Glu Leu Trp Leu Val Met Lys Leu Leu
 100 105 110

Ser Gly Gly Ser Met Leu Asp Ile Ile Lys Tyr Ile Val Asn Arg Gly
 115 120 125

Glu His Lys Asn Gly Val Leu Glu Glu Ala Ile Ile Ala Thr Ile Leu
 130 135 140

Lys Glu Val Leu Glu Gly Leu Asp Tyr Leu His Arg Asn Gly Gln Ile
 145 150 155 160

His Arg Asp Leu Lys Ala Gly Asn Ile Leu Leu Gly Glu Asp Gly Ser
 165 170 175

Val Gln Ile Ala Asp Phe Gly Val Ser Ala Phe Leu Ala Thr Gly Gly
180 185 190

Asp Val Thr Arg Asn Lys Val Arg Lys Thr Phe Val Gly Thr Pro Cys
195 200 205

Trp Met Ala Pro Glu Val Met Glu Gln Val Arg Gly Tyr Asp Phe Lys
210 215 220

Ala Asp Met Trp Ser Phe Gly Ile Thr Ala Ile Glu Leu Ala Thr Gly
225 230 235 240

Ala Ala Pro Tyr His Lys Tyr Pro Pro Met Lys Val Leu Met Leu Thr
245 250 255

Leu Gln Asn Asp Pro Pro Thr Leu Glu Thr Gly Val Glu Asp Lys Glu
260 265 270

Met Met Lys Lys Tyr Gly Lys Ser Phe Arg Lys Leu Leu Ser Leu Cys
275 280 285

Leu Gln Lys Asp Pro Ser Lys Arg Pro Thr Ala Ala Glu Leu Leu Lys
290 295 300

Cys Lys Phe Phe Gln Lys Ala Lys Asn Arg Glu Tyr Leu Ile Glu Lys
305 310 315 320

Leu Leu Thr Arg Thr Pro Asp Ile Ala Gln Arg Ala Lys Lys Val Arg
325 330 335

Arg Val Pro Gly Ser Ser Gly His Leu His Lys Thr Glu Asp Gly Asp
340 345 350

Trp Glu Trp Ser Asp Asp Glu Met Asp Glu Lys Ser Glu Glu Gly Lys
355 360 365

Ala Ala Phe Ser Gln Glu Lys Ser Arg Arg Val Lys Glu Glu Asn Pro
370 375 380

Glu Ile Ala Val Ser Ala Ser Thr Ile Pro Glu Gln Ile Gln Ser Leu
385 390 395 400

Ser Val His Asp Ser Gln Gly Pro Pro Asn Ala Asn Glu Asp Tyr Arg
405 410 415

Glu Ala Ser Ser Cys Ala Val Asn Leu Val Leu Arg Leu Arg Asn Ser
420 425 430

Arg Lys Glu Leu Asn Asp Ile Arg Phe Glu Phe Thr Pro Gly Arg Asp
435 440 445

Thr Ala Asp Gly Val Ser Gln Glu Leu Phe Ser Ala Gly Leu Val Asp
450 455 460

Gly His Asp Val Val Ile Val Ala Ala Asn Leu Gln Lys Ile Val Asp

465	470	475	480
-----	-----	-----	-----

Asp Pro Lys Ala Leu Lys Thr Leu Thr Phe Lys Leu Ala Ser Gly Cys			
485		490	495

Asp Gly Ser Glu Ile Pro Asp Glu Val Lys Leu Ile Gly Phe Ala Gln			
500		505	510

Leu Ser Val Ser			
515			

<210> 7

<211> 414

<212> PRT

<213> Mammalian (Human) STLK4

<220>

<223> "Xaa" stands for any amino acid.

<400> 7

Lys Ser Gly Val Leu Asp Xaa Ser Thr Ile Ala Thr Ile Leu Arg Glu			
1	5	10	15

Val Leu Glu Gly Leu Glu Tyr Leu His Lys Xaa Gly Gln Ile His Arg			
20		25	30

Asp Val Lys Ala Gly Asn Ile Leu Xaa Gly Glu Asp Gly Ser Val Gln			
35	40		45

Ile Ala Asp Phe Gly Val Ser Ala Phe Leu Ala Thr Gly Gly Asp Ile			
50		55	60

Thr Arg Asn Lys Val Arg Lys Thr Phe Val Gly Thr Pro Cys Trp Met			
65	70	75	80

Ala Pro Glu Val Met Glu Gln Val Arg Gly Tyr Asp Phe Lys Ala Asp			
85		90	95

Ile Trp Ser Phe Gly Ile Thr Ala Ile Glu Leu Ala Thr Gly Ala Ala			
100		105	110

Pro Tyr His Lys Tyr Pro Pro Met Lys Val Leu Met Leu Thr Leu Gln			
115	120		125

Asn Asp Pro Pro Ser Leu Glu Thr Gly Val Gln Asp Lys Glu Met Leu			
130	135		140

Lys Lys Tyr Gly Lys Ser Phe Arg Lys Met Ile Ser Leu Cys Leu Gln			
145	150	155	160

Lys Asp Pro Glu Lys Arg Pro Thr Ala Ala Glu Leu Leu Arg His Lys			
165		170	175

10

Phe Phe Gln Lys Ala Lys Asn Lys Glu Phe Leu Gln Glu Lys Thr Leu
180 185 190

Gln Arg Ala Pro Thr Ile Ser Glu Arg Ala Lys Lys Val Arg Arg Val
195 200 205

Pro Gly Ser Ser Gly Arg Leu His Lys Thr Glu Asp Gly Gly Trp Glu
210 215 220

Trp Ser Asp Asp Glu Phe Asp Glu Glu Ser Glu Glu Gly Lys Ala Ala
225 230 235 240

Ile Ser Gln Leu Arg Ser Pro Arg Val Lys Glu Ser Ile Ser Asn Ser
245 250 255

Glu Leu Phe Pro Thr Thr Asp Pro Val Gly Thr Leu Leu Gln Val Pro
260 265 270

Glu Gln Ile Ser Ala His Leu Pro Gln Pro Ala Gly Gln Ile Ala Thr
275 280 285

Gln Pro Thr Gln Val Ser Leu Pro Pro Thr Ala Glu Pro Ala Lys Thr
290 295 300

Ala Gln Ala Leu Ser Ser Gly Ser Gly Ser Gln Glu Thr Lys Ile Pro
305 310 315 320

Ile Ser Leu Val Leu Arg Leu Arg Asn Ser Lys Lys Glu Leu Asn Asp
325 330 335

Ile Arg Phe Glu Phe Thr Pro Gly Arg Asp Thr Ala Glu Gly Val Ser
340 345 350

Gln Glu Leu Ile Ser Ala Gly Leu Val Asp Gly Arg Asp Leu Val Ile
355 360 365

Val Ala Ala Asn Leu Gln Lys Ile Val Glu Glu Pro Gln Ser Asn Arg
370 375 380

Ser Val Thr Phe Lys Leu Ala Ser Gly Val Glu Gly Ser Asp Ile Pro
385 390 395 400

Asp Asp Gly Lys Leu Ile Gly Phe Ala Gln Leu Ser Ile Ser
405 410

<210> 8
<211> 274
<212> PRT
<213> Mammalian (Human) STLK5

<400> 8

Leu Ile Cys Thr His Phe Met Asp Gly Met Asn Glu Leu Ala Ile Ala
1 5 10 15

Tyr Ile Leu Gln Gly Val Leu Lys Ala Leu Asp Tyr Ile His His Met
 20 25 30

Gly Tyr Val His Arg Ser Val Lys Ala Ser His Ile Leu Ile Ser Val
 35 40 45

Asp Gly Lys Val Tyr Leu Ser Gly Leu Arg Ser Asn Leu Ser Met Ile
 50 55 60

Ser His Gly Gln Arg Gln Arg Val Val His Asp Phe Pro Lys Tyr Ser
 65 70 75 80

Val Lys Val Leu Pro Trp Leu Ser Pro Glu Val Leu Gln Gln Asn Leu
 85 90 95

Gln Gly Tyr Asp Ala Lys Ser Asp Ile Tyr Ser Val Gly Ile Thr Ala
 100 105 110

Cys Glu Leu Ala Asn Gly His Val Pro Phe Lys Asp Met Pro Ala Thr
 115 120 125

Gln Met Leu Leu Glu Lys Leu Asn Gly Thr Val Pro Cys Leu Leu Asp
 130 135 140

Thr Ser Thr Ile Pro Ala Glu Glu Leu Thr Met Ser Pro Ser Arg Ser
 145 150 155 160

Val Ala Asn Ser Gly Leu Ser Asp Ser Leu Thr Thr Ser Thr Pro Arg
 165 170 175

Pro Ser Asn Gly Asp Ser Pro Ser His Pro Tyr His Arg Thr Phe Ser
 180 185 190

Pro His Phe His His Phe Val Glu Gln Cys Leu Gln Arg Asn Pro Asp
 195 200 205

Ala Arg Pro Ser Ala Ser Thr Leu Leu Asn His Ser Phe Phe Lys Gln
 210 215 220

Ile Lys Arg Arg Ala Ser Glu Ala Leu Pro Glu Leu Leu Arg Pro Val
 225 230 235 240

Thr Pro Ile Thr Asn Phe Glu Gly Ser Gln Ser Gln Asp His Ser Gly
 245 250 255

Ile Phe Gly Leu Val Thr Asn Leu Glu Glu Leu Glu Val Asp Asp Trp
 260 265 270

Glu Phe

<210> 9
<211> 3798
<212> DNA

<213> Mammalian (Human) ZC1

<400> 9

gagaccatgg cgaacgactc tcccgcaaa agtctggtgg acatcgacct	ctccctccctg	60
cgggatcctg ctggatttt tgagctggtg gaagtgggt gaaaatggcac	ctatggacaa	120
gtctataagg gtcgacatgt taaaacgggt cagttggcag ccataaagt	tatggatgtc	180
actgaggatg aagaggaaga aatcaaactg gagataaata tgctaaagaa	atactctcat	240
cacagaaaca ttgcaacata ttatggtgct ttcatcaaaa agagccctcc	aggacatgt	300
gaccaactct ggcttgttat ggagttctgt ggggctgggt ccattacaga	ccttgtgaag	360
aacaccaaag ggaacacacact caaagaagac tggatcgctt acatctccag	agaaaatccctg	420
aggggactgg cacatcttca cattcatcat gtgattcacc gggatatcaa	gggccagaat	480
gtgttgcgtga ctgagaatgc agaggtaaa cttgttgact ttgtgtgag	tgctcaactg	540
gacaggactg tggggcggag aaatacgttc ataggcactc cctactggat	ggctccctgag	600
gtcatcgct gtgatgagaa cccagatgcc acctatgatt acagaagtga	tctttggct	660
tgtggcatta cagccattga gatggcagaa ggtgctcccc ctctctgtga	catgcattca	720
atgagagcac tttttctcat tcccagaaac cttccctcccc ggctgaagtc	aaaaaaatgg	780
tcgaagaagt ttttagttt tatagaaggg tgcctggta agaattacat	gcagcggccc	840
tctacagagc agttttgaa acatccttt ataaggatc agccaaatga	aaggcaagtt	900
agaatccagc ttaaggatca tatagatcgt accaggaaga agagaggcga	gaaagatgaa	960
actgagatg agtacagtgg gagtgaggaa gaagaggagg aagtgcctga	acaggaagga	1020
gagccaagtt ccattgtgaa cgtgcctggt gagtctactc ttgcggcaga	tttcctgaga	1080
ctgcagcagg agaacaagga acgttccgag gctttcggg gacaacagtt	actacaggag	1140
caacagctcc gggagcagga agaatataaa aggcaactgc tggcagagag	acagaagcgg	1200
attgagcagc agaaagaaca gaggcgcacgg cttagaagagc aacaaggag	agagcgggaa	1260
gctagaaggc agcaggaacg tgaacagcga aggagagaaac aagaagaaaa	gaggcgtcta	1320
gaggagttgg agagaaggcg caaagaagaa gaggagagga gacgggcaga	agaagaaaaag	1380
aggagagttt aaagagaaca ggatataatc aggccagacg tagaagagga	gcagcggcac	1440
tttggagtcc ttccagcagca gctgcctccag gaggcggca ttttacttgg	gtgccgatgg	1500
cgggagatgg aggagcaccc gcaggcagag aggctccaga ggcagttgca	acaagaacaa	1560
gcataatctcc ttttcttaca gcatgaccat aggaggccgc acccgagca	ctcgccagcag	1620
ccggccaccac cgccaggcagga aaggagcaag ccaagctcc atgtcccgaa	gcccaaagcc	1680
caactacgagc ctgctgaccc agcgcgagag gtggaaagata gatttagaa	aactaaccac	1740
agctcccctg aagcccgatc taagcagaca ggcagatgt tggagccacc	agtgccttcc	1800
cgatcagagt tttttccaa tggcaactcc gagtctgtgc atcccgccct	gcagagacca	1860
cgccgagccac aggttccctgt gagaacaaca ttcgcctccc ctgttctgtc	ccgtcgagat	1920
tccccactgc agggcagttgg gcagcagaat agccaggcag gacagagaaa	ctccaccagt	1980
attgagccca ggcttctgtg ggagagagtg gagaagctgg tgcccgacc	tggcagttggc	2040
agctccctcg ggtccagcaa ctcaggatcc cagcccggtt ctcaccctgg	gtctcagagt	2100
ggctccgggg aacgccttcag agttagatca tcataaagt ctgaaggctc	tccatctcg	2160
cgccctggaaa atgcagttgaa aaaacctgaa gataaaaaagg aagtttcag	acccctcaag	2220
cctgctgatc tgaccgcact ggccaaagag cttcgagcag tggagatgt	acggccaccc	2280
cacaaagtaa cggactactc ctcatccagt gaggagtcgg ggacgacgga	tgaggaggac	2340
gacgatgtgg agcaggaagg ggctgcacgg tccacccctcg qaccaqagga	caccagagca	2400
cggtcatctc tgaattttag caatggtgg acggaatctg tgaaaaccat	gattgtccat	2460
gatgatgttag aagttagcc ggccatgacc ccatccaagg agggactct	aatcgccgc	2520
cggtactcgat cgcgttagtag cacactccag aaacacaaat cttccctcc	ctttacaccc	2580
tttatagacc ccagattact acagattct ccatctagcg gaacaacagt	gacatctgtg	2640
gtgggatttt cctgtgtatgg gatgagacca gaagccataa ggcaagatcc	tacccggaaa	2700
ggctcagtgg tcaatgtgaa tcctaccaac actaggccac agagtacac	cccgagatt	2760
cgtaaataca agaagaggtt taactcttag attctgtgtg ctgcctttag	gggagtgaat	2820
ttgcttagtgg gtacagagag tggcctgtatg ctgctggaca gaagtggcca	agggaaaggtc	2880
tatcccttta tcaaccgaag acgattcaa caaatggacg tacttgaggg	tttgaatgtc	2940
ttgggtgacaa tatctggcaa aaaggataag ttacgtgtct actatttgc	ctggtaaga	3000
aataaaaatac ttccacaatga tccagaagtt gagaagaagc agggatggac	aaccgttaggg	3060
gatggaaag gatgtgtaca ttataaagtt gtaaaaatatg aaagaatcaa	atttctggtg	3120

atggcctta	agtcatgg	agaattgta	cataagccat	tactggta	tctca	tcacaaat	3180
gaggaaggcc	agaggtgaa	agtcatat	ggatctgt	ctgattcca	tgctgtt	gat	3240
gtggattcag	gtacatcta	tgacattat	ctaccaacac	atatccagt	tagcatcaa	ccc	3300
ccccatgcaa	tcatcatcct	ccccaaatac	gatggatgg	agcttctgg	gtgctatgaa	at	3360
gatgaggggg	tttatgtaaa	cacatatgga	aggatcacca	aggatgttagt	tctacatgtgg	gg	3420
ggagagatgc	ctacatcagt	agcatatatt	cgatccaatc	agacaatggg	ctggggagag	aagg	3480
aaggccatag	agatccgatc	tgtggaaact	ggtca	atgtgtgtt	catgcacaaa	aggg	3540
agggctcaaa	gactaaaatt	cttgcgtgaa	cgcaatgaca	aggtgttctt	tgcctctgtt	cggt	3600
cggtctgg	gcagcagtca	ggttatttc	atgaccttag	gcaggacttc	tcttctgac	ttgt	3660
ttgtagaagc	agtgtgatcc	aggattact	ggcctccaga	gtctcaaga	tctgagaac	ttgg	3720
ttggaaattcc	ttgtact						3780
							3798

<210> 10
<211> 4055
<212> DNA
<213> Mammalian (Human) ZC2

<400> 10

gttttcgggg	aggcttatga	gggtcgcat	gtcaaaacgg	gccagcttgc	agccatcaag	60
gttatggatg	tcacagggaa	tgaagaggaa	gaaatcaa	aagaaattaa	catgttgaag	120
aaatattctc	atcacccgaa	tattgtaca	tactatgg	cttttatcaa	aaagaaccca	180
ccaggcatgg	atgaccaact	ttgggtgg	atggagttt	gtggtgc	ctctgtcacc	240
gacctgatca	agaacacaaa	agtaaacacq	ttgaaagagg	atggattgc	atacatctgc	300
agggaaatct	tacggggct	gagtccat	caccagcata	aagtattca	tcgagatatt	360
aaaggcaaa	atgtctgct	gactaaaaat	gcagaagttt	aactatgtgg	cttggagtc	420
agtgc	ttgatcgaac	atgggcagg	aggaataactt	tcattgg	tcctactgg	480
atggcaccag	aagtattgc	ctgtatgaa	acccagatg	ccacatatga	tttcaagagt	540
gacttgggt	cttgggtat	caccgcatt	gaaatggcag	aagtgc	ccctctctgt	600
gacatgcacc	ccatgagac	tcttcctc	atccccccg	atccagcgcc	tcggctgaag	660
tctaagaat	ggtcaaaaaa	attccagtca	tttattgaga	gctgttgg	aaagaatcac	720
agccagcgac	cagcaacaga	acaattgtat	aagcatccat	ttatacgaga	ccaaacctaat	780
gagcgacagg	tccgattca	actcaaggac	catattgata	gaacaaagaa	gaagcgagga	840
aaaaaagatg	agacagatg	tgagtacat	ggaagtgg	aagaagagga	ggagaatgac	900
tcaggagac	ccagctccat	cctgaatctg	ccaaggagt	cgacgctgcg	gagggactt	960
ctgaggctgc	agctggccaa	caaggagcgt	tctgaggccc	taaggaggca	gcagctggag	1020
cagcagcagc	gggagaatga	ggagcacaag	cgccagctgc	tggccgagcg	tcagaagcg	1080
atcgaggagc	agaaaagagca	gaggcggcgg	ctggaggagc	aacaaaggcg	agagaaggag	1140
ctgcggaa	agcaggagag	ggagcagcgc	cggcactatg	aggagcagat	gcgcggggag	1200
gaggagagga	ggcgtcg	gcatgaacag	gaatataagc	gcaacaatt	ggaagaacag	1260
agacaagcag	aaagactgca	gaggcagct	aagcaagaaa	gagactactt	agttccctt	1320
cagcatcagc	ggcaggagca	gaggcgt	gagaagaagc	cactgtacca	ttacaagaa	1380
ggaatgatc	ctagtgagaa	gccagcatgg	gccaaggagg	taqaagaacg	gtcaaggctc	1440
aaccggcaaa	gttcccctgc	catgcctcac	aagggtgcca	acaggatatc	tgaccccaac	1500
ctgccccca	ggtcggagtc	cttcagcatt	agtggagttc	agctgctcg	aacacccccc	1560
atgc	cagtc	ccagatccca	catctgg	ctgtaaaatc	ccagggac	1620
gccttgaccg	cctcccagtc	agtgcacgag	cagcccacaa	aggcctc	tggtttgc	1680
gaggctotga	acgtgacctc	ccaccgcgt	gagatgccac	gccagaactc	agatcccacc	1740
tcggaaaatc	ctcctctccc	cactcgcatt	gaaaagttt	accgaagctc	ttgggtacga	1800
caggaagaag	acattccacc	aaaggtgcct	caaagaacaa	cttctatatc	cccagcatta	1860
gccagaaaga	attctccctgg	gaatggtagt	gctctggac	ccagactagg	atctcaaccc	1920
atcagagcaa	gcaaccctga	tctccggaga	actgagccca	tctggagag	cccctgcag	1980
aggaccagca	gtggcagttc	ctccagctcc	agcaccccta	gctcccagcc	cagctccaa	2040

ggaggctccc	agcctggate	acaagcagga	tccagtgaac	gcaccagagt	tcgagccaac	2100
agtaagtca	aaggatcacc	tgtgctcccc	catgagcctg	ccaaggtgaa	accagaagaa	2160
tccaggaca	ttacccggcc	cagtcgacca	gctagctaca	aaaaagctat	agatgaggat	2220
ctgacggcat	tagccaaaga	actaagagaa	ctccggattg	aagaacaaa	ccgccaatg	2280
aagaaggta	ctgattactc	ctcctccagt	gaggagtca	aaagtagcga	ggaagaggag	2340
gaagatggag	agagcgagac	ccatgatggg	acagtggctg	tcagcgacat	acccagactg	2400
ataccaacag	gagctccagg	cagcaacgag	cagtacaatg	tggaaatggt	ggggacgcat	2460
gggctggaga	cctctcatgc	ggacagttc	agcggcagta	tttcaagaga	aggaaccttg	2520
atgattagag	agacgtctgg	agagaagaag	cgatctggcc	acagtgcac	caatggctt	2580
gctggccaca	tcaacctccc	tgacctggt	cagcagagcc	attctccagc	tggaaaccccg	2640
actgagggac	tggggchggt	ctcaacccat	tcccaggaga	tggactctgg	gactgaatat	2700
ggcatgggga	gcagcaccaa	agcctcttc	acccccttg	tggaccccg	agtataccag	2760
acgtctccca	ctgatgaaga	tgaagaggat	gaggaatcat	cagccgcagc	tctgtttact	2820
ggcgaacttc	ttaggcaaga	acaggccaa	ctcaatgaag	caagaaagat	tccgtggta	2880
aatgtaaacc	caaccaacat	tcggcctcat	agcgacacac	cagaaatcag	aaaatacaag	2940
aaacgattca	actcagaaat	acttttgca	gctctgtggg	gtgtaaacct	tctggtgggg	3000
actgaaaatg	gcctgtatct	tttgaccga	agtggcaag	gcaaagtcta	taatctgatc	3060
aaccggaggc	gatttcagca	gatggatgt	ctagagggac	tgaatgtcct	tgtgacaatt	3120
tcagggaaaga	agaataagct	acgagttac	tatcttcat	ggttaagaaa	cagaatacta	3180
cataatgacc	cagaagttaga	aaagaaacaa	ggctggatca	ctgttggga	cttggaaaggc	3240
tgtatacatt	ataaaagtgt	taaatatgaa	aggatcaa	ttttggat	tgccttaaag	3300
aatgctgtgg	aaatataatgc	ttgggctct	aaaccgtatc	ataaaattcat	ggcatttaag	3360
tctttgcag	atctccagca	caagcctctg	ctagttgatc	tcacggtaga	agaaggtaa	3420
agattaaagg	ttattttg	ttcacacact	ggtttccatg	taattgtatgt	tgattcaagga	3480
aactcttatg	atatctacac	accatctcat	attcagggca	atatcactcc	tcatgttatt	3540
gtcatcttgc	ctaaaacaga	tggatggaa	atgcttgtt	gctatgagga	tgagggggtg	3600
tatgtaaaca	cctatggccg	gataactaag	gatgtggc	tccaatgggg	agaaatgccc	3660
acgtctgtgg	cctacattca	ttccaatcag	ataatggct	ggggcgagaa	agctattgag	3720
atccggctag	tggaaacagg	acatttggat	ggagtattta	tgcataagcg	agctcaaagg	3780
ttaaagttc	tatgtgaaag	aatgataag	gtatttttg	catccgtgcg	atctggagga	3840
agtagccaag	tgtttttcat	gaccctcaac	agaaattcca	tgtgaactg	gtaacagaag	3900
agcacttggc	acttatcttc	atggcgat	ttctaaat	aaagaacata	actcatgtgg	3960
acttatgcca	gtctagaggc	agaatcagaa	ggcttgggtt	aacatatcgc	tttcccttt	4020
tcctctccct	ccgcccctcc	cagtacagtc	catct			4055

<210> 11
 <211> 4133
 <212> DNA
 <213> Mammalian (Human) ZC3

<400> 11

gcattttgggg	aggtgtatga	gggtcgcat	gtcaagacgg	ggcagctggc	tgccatcaag	60
gtcatggatg	tcacggagga	cgaggaggaa	gagatcaa	aggagatcaa	catgctaaa	120
aagtactctc	accaccgcaa	catgccacc	tactacggag	ctttcatcaa	gaagagcccc	180
ccggaaacg	atgaccagct	ctggctggt	atggagttt	gtgggtctgg	ttcagtact	240
gacctggtaa	agaacacaaa	aggcaacgccc	ctgaaggagg	actgtatcgc	ctatatctgc	300
agggagatcc	tcaggggtct	ggcccatctc	catgcccaca	aggtgatcca	tcgagacatc	360
aaggggcaga	atgtgctgt	gacagagaat	gctgagggtca	agctagtgg	ttttgggttg	420
agtgcgtcgc	tggaccgcac	cgtggcaga	cggaacactt	tcatggac	tccctactgg	480
atggctccag	aggtcatcgc	ctgtgatgag	aaccctgatg	ccacctatga	ttacaggagt	540
gatatttggt	ctcttagaat	cacagccatc	gagatggcag	aggagcccc	ccctctgtgt	600
gacatgcacc	ccatgcgagc	cctttctcc	attcctcgga	accctccgccc	caggctcaag	660
tccaagaagt	ggtctaagaa	gttcattgac	ttcattgaca	catgtctcat	caagacttac	720

ctgagccgcc cacccacgga gcagctactg aagttccct tcatccggg ccagcccacg	780
gagcggcagg tccgcattcca gcttaaggac cacattgacc gatcccgaa gaagccccgt	840
gagaaaaggagg agacagaata tgagtacagc ggcagcgagg aggaagatga cagccatgga	900
gaggaaggagg agccaagctc catcatgaac gtgcctggag agtcgactct acgccccggag	960
tttctccggc tccagcagga aaataagagc aactcagagg cttaaaaaca gcagcagcag	1020
ctgcagcagc agcagcagcg agaccccgag gcacacatca aacacctgct gcaccagcgg	1080
cagcggcgca tagaggagca gaaggaggag cggcggcccg tgagggagca acagccccgg	1140
gagcggggcgc agcggaaagct gcagggagaag gagcagcagc ggccggctgga ggacatgcag	1200
gctctgcggc gggaggaggag gcggcggcag gcggagcgcg agcaggaata tattcgtcac	1260
aggctagagg aggagcagcg acagctcgag atccttcagc aacagctgct ccaggaacag	1320
gccctgtgc tgaataacaa gcggaaagcag ctggaggagc agcggcagtc agaacgtctc	1380
cagaggcagc tgcaagcagga gcatgcctac ctcaagttccc tgcaagcagca gcaacacgcag	1440
cagcagcttc agaaacaaca gcagcagcag ctccctgcgt gggacaggaa gccctgtac	1500
cattatggtc gggcatgaa tcccgtgac aaaccaggct gggcccgaga ggtagaagag	1560
agaacaagga tgaacaagca gcagaactct cccttggca agaccaagcc aggcagcagc	1620
gggcctgagc ccccccattccc ccaggcctcc ccagggccccc caggaccctt ttcccagact	1680
cctcttatgc agaggccgt ggagcccccag gagggaccgc acaagagcct ggtggcacac	1740
cgggtcccac tgaagccata tgcaagcacct gtaccccgat cccagtcct gcaaggaccag	1800
cccaccccaa acttggctgc ttcccagcc tcccatgacc ccgaccctgc catccccca	1860
cccactgcca cggccagtgcc cggaggagct gtcatccgca agaattcaga ccccacctct	1920
gaaggacctg gccccagccc gaatccccca gcctgggtcc gcccagataa cgagggccca	1980
cccaaggtgc ctcagaggac ctcatctatc gcaactgccc ttaacaccag tggggccgga	2040
gggtcccgcc cagcccaggc agtccgtgcc agacctcgca gcaactccgc ctggcaaattc	2100
tatctgcaaa ggcgggcaga gcggggcacc ccaaagcctc cagggccccc tgctcagccc	2160
cctggcccgcc ccaacgcctc tagtaacccc gacctcagga ggagcgcaccc tggctggaa	2220
cgctcggaca gcgtccctcc agcctctac gggcacctcc cccaggctgg ctcaactggag	2280
cggAACCGCG tggaggtctc ctccaaaccg gacagctccc ctgtgctctc ccctggaaat	2340
aaagccaaagc ccgacgacca ccgctcacgg ccagggccgc ccgcagactt tgtgttgctg	2400
aaagagcggc ctctggacga ggcccctcgg cctcccaaga aggcacatgga ctactcgctg	2460
tccagcggagg aggtggaaag cagtgaggac gacgaggagg aagcgaagg cggccagca	2520
gaggggagca gagataccccc tggggccgc gatggggata cagacagcgt cagcaccatg	2580
gtggtccacg acgtcgagga gatcacccgg acccagcccc catacggggg cggcaccatg	2640
gtggtccagc gcacccctga agaggagcgg aacctgctgc atgctgacag caatgggtac	2700
acaaacctgc ctgacgtggt ccagcccagc cactcacca ccgagaacag caaaggccaa	2760
agcccacccct cgaaggatgg gagtgggtac taccagtcgtc gtgggcttgtt aaaggcccct	2820
ggcaagagct cgttcacgt gtttggat cttagggatct accagcctgg aggcaagtgg	2880
gacagcatcc ccatcacagc cctagtggtt ggagaggca ctccgctcga ccagctcag	2940
taegacgtga ggaagggttc tgggtcaac gtgaatccca ccaaaccctg gcccacagt	3000
gagaccctg agatccggaa gtacaagaag cgattcaact ccgagatctt ctgtcagcc	3060
ctttgggggg tcaacctgtc ggtggcacc gagaacggc tgatgttgct ggaccgaagt	3120
gggcaggggca aggtgtatgg actcattggg cggcgcacgt tccagcagat ggatgtgctg	3180
gaggggctca acctgctcat caccatcta gggaaaagga acaaactgcg ggtgttattac	3240
ttgtcctggc tccggaaacaa gattctgcac aatgacccag aagtggagaa gaagcagggc	3300
tggaccaccg tgggggacat ggaggcgtc gggcactacc gtgttgtaa atacgagcgg	3360
attaagtcc tggtcategc cctcaagagc tccgtggagg tggatgtcgt ggcacccaaa	3420
ccctaccaca aattcatggc cttcaagtcc tttggccgacc tccccccaccc ccctctgtc	3480
gtcgacccctga cagtagagga gggcaggcgg ctcaaggctca tctatggctc cagtgctggc	3540
ttccatgtctg tggatgtcga ctgggggaaac agctatgaca tctacatccc tggcacatc	3600
cagagccaga tcaacgccccca tgccatcatc ttccctccca acaccgcacgg catggagatg	3660
ctgctgtgtc acgaggacga ggggtctac gtcaacacgt acggcgcacat cattaaggat	3720
gtggtgcgtc agtgggggg gatgcctact tctgtggctt acatctgtctc caaccagata	3780
atgggctggg gtgagaaagc cattgagatc cgctctgtgg agacggggca cctcgacggg	3840
gtcttcatgc acaaacgac tcagaggctc aagttccctgt gtgagcggaa tgacaagggt	3900
ttttttgcct cagtcgcct tggggcaggc agccaagttt acttcatgac tctgaaccgt	3960
aaccgcacca tgaactgggt acggggccctt gggctggggc tggccacac tggacccagc	4020
tctccctctg cagccaggct tccggggcccg ccccttttc ccctccctgg gctttgctt	4080

ttactggttt gatttcactg gagcctgctg ggaacgtgac ctctgacccc tga

4133

<210> 12
<211> 1459
<212> DNA
<213> Mammalian (Human) ZC4

<400> 12

caatgttaac ccactctatg tctctcctgc atgtaaaaaa ccactaatcc acatgttatga	60
aaaggagttc acttctgaga tctgctgtgg ttctttgtgg ggagtcaatt tgctgttggg	120
aacccgatct aatctatatac tgatggacag aagtggaaag gctgacatta ctaaacttat	180
aaggcgaaga ccattccgcc agattcaagt cttagagcca ctcaatttgc tgattaccat	240
ctcaggtcat aagaacagac ttcgggtgta tcatctgacc tggttgagga acaagattt	300
gaataatgat ccagaaagta aaagaaggca agaagaatg ctgaagacag aggaagcctg	360
caaagctatt gataagttaa caggctgtga acacttcagt gtccccaac atgaagaaac	420
aacatatatt gcaattgctt tgaaatcatc aattcacett tatgcattggg caccaaagtc	480
ctttgatgaa agcactgcta taaaagtatt tccaacacett gatcataagc cagtgacagt	540
tgacctggct attggttctg aaaaaagact aaagattttc ttcagctcag cagatggata	600
tcacccatc gatgcagaat ctgaggttat gtctgatgtg accctgccaa agaatcccct	660
ggaaatcatt ataccacaga atatcatcat tttacctgat tgcttggaa ttggcatgat	720
gctcacccctc aatgctgaag ccctctctgt ggaagcaat gaacaactct tcaagaagat	780
ccttgaaatg tgaaaagaca taccatctc tatagtttt gaatgtacac agcgaaccac	840
aggatgggc caaaaggcca ttgaagtgcg ctcttgcaa tccagggttc tggaaagtga	900
gctgaagcgc aggtcaatta agaagctgag attcctgtgc acccggttgc acaagctgtt	960
ctttacctct accctgcgc atcaccacag ccgggttac ttcatgacac ttggaaaact	1020
tgaagagctc caaagcaatt atgatgtcta aaagtttcca gtgatttatt accacattat	1080
aaacatcatg tataggcagt ctgcatttc agattcaga gattaaatga gtattcagtt	1140
ttatTTTtag taaaagattaa atccaaaact ttactttaa tgtagcacag aatagttta	1200
atgagaaatg cagcttatg tataaaatta actatagcaa gctctaggtt ctccaatgg	1260
gtacaatgtc tttgcacaa actttgtaac ttttgttact gtgaattcaa acattactct	1320
ttggacagtt tggacagtt ctgtatttcg attttacaac atggagtaaa gaaacctgtt	1380
atgaattaga ttacaaggcag cttcaaaaag aattggcaact gggataagat tttcagaaa	1440
agaaaaacat cggcaaact	1459

<210> 13
<211> 1239
<212> PRT
<213> Mammalian (Human) ZC1

<400> 13

Met Ala Asn Asp Ser Pro Ala Lys Ser Leu Val Asp Ile Asp Leu Ser			
1	5	10	15

Ser Leu Arg Asp Pro Ala Gly Ile Phe Glu Leu Val Glu Val Val Gly			
20	25	30	

Asn Gly Thr Tyr Gly Gln Val Tyr Lys Gly Arg His Val Lys Thr Gly

35

40

45

Gln Leu Ala Ala Ile Lys Val Met Asp Val Thr Glu Asp Glu Glu Glu
 50 55 60

Glu Ile Lys Leu Glu Ile Asn Met Leu Lys Lys Tyr Ser His His Arg
 65 70 75 80

Asn Ile Ala Thr Tyr Tyr Gly Ala Phe Ile Lys Lys Ser Pro Pro Gly
 85 90 95

His Asp Asp Gln Leu Trp Leu Val Met Glu Phe Cys Gly Ala Gly Ser
 100 105 110

Ile Thr Asp Leu Val Lys Asn Thr Lys Gly Asn Thr Leu Lys Glu Asp
 115 120 125

Trp Ile Ala Tyr Ile Ser Arg Glu Ile Leu Arg Gly Leu Ala His Leu
 130 135 140

His Ile His His Val Ile His Arg Asp Ile Lys Gly Gln Asn Val Leu
 145 150 155 160

Leu Thr Glu Asn Ala Glu Val Lys Leu Val Asp Phe Gly Val Ser Ala
 165 170 175

Gln Leu Asp Arg Thr Val Gly Arg Arg Asn Thr Phe Ile Gly Thr Pro
 180 185 190

Tyr Trp Met Ala Pro Glu Val Ile Ala Cys Asp Glu Asn Pro Asp Ala
 195 200 205

Thr Tyr Asp Tyr Arg Ser Asp Leu Trp Ser Cys Gly Ile Thr Ala Ile
 210 215 220

Glu Met Ala Glu Gly Ala Pro Pro Leu Cys Asp Met His Pro Met Arg
 225 230 235 240

Ala Leu Phe Leu Ile Pro Arg Asn Pro Pro Pro Arg Leu Lys Ser Lys
 245 250 255

Lys Trp Ser Lys Lys Phe Phe Ser Phe Ile Glu Gly Cys Leu Val Lys
 260 265 270

Asn Tyr Met Gln Arg Pro Ser Thr Glu Gln Leu Leu Lys His Pro Phe
 275 280 285

Ile Arg Asp Gln Pro Asn Glu Arg Gln Val Arg Ile Gln Leu Lys Asp
 290 295 300

His Ile Asp Arg Thr Arg Lys Lys Arg Gly Glu Lys Asp Glu Thr Glu
 305 310 315 320

Tyr Glu Tyr Ser Gly Ser Glu Glu Glu Glu Glu Val Pro Glu Gln
 325 330 335

Glu Gly Glu Pro Ser Ser Ile Val Asn Val Pro Gly Glu Ser Thr Leu
340 345 350

Arg Arg Asp Phe Leu Arg Leu Gln Gln Glu Asn Lys Glu Arg Ser Glu
355 360 365

Ala Leu Arg Arg Gln Gln Leu Leu Gln Glu Gln Leu Arg Glu Gln
370 375 380

Glu Glu Tyr Lys Arg Gln Leu Leu Ala Glu Arg Gln Lys Arg Ile Glu
385 390 395 400

Gln Gln Lys Glu Gln Arg Arg Leu Glu Glu Gln Gln Arg Arg Glu
405 410 415

Arg Glu Ala Arg Arg Gln Gln Glu Arg Glu Gln Arg Arg Glu Gln
420 425 430

Glu Glu Lys Arg Arg Leu Glu Glu Leu Glu Arg Arg Arg Lys Glu Glu
435 440 445

Glu Glu Arg Arg Ala Glu Glu Glu Lys Arg Arg Val Glu Arg Glu
450 455 460

Gln Glu Tyr Ile Arg Arg Gln Leu Glu Glu Glu Gln Arg His Leu Glu
465 470 475 480

Val Leu Gln Gln Leu Leu Gln Glu Gln Ala Met Leu Leu Glu Cys
485 490 495

Arg Trp Arg Glu Met Glu Glu His Arg Gln Ala Glu Arg Leu Gln Arg
500 505 510

Gln Leu Gln Gln Glu Gln Ala Tyr Leu Leu Ser Leu Gln His Asp His
515 520 525

Arg Arg Pro His Pro Gln His Ser Gln Gln Pro Pro Pro Pro Gln Gln
530 535 540

Glu Arg Ser Lys Pro Ser Phe His Ala Pro Glu Pro Lys Ala His Tyr
545 550 555 560

Glu Pro Ala Asp Arg Ala Arg Glu Val Glu Asp Arg Phe Arg Lys Thr
565 570 575

Asn His Ser Ser Pro Glu Ala Gln Ser Lys Gln Thr Gly Arg Val Leu
580 585 590

Glu Pro Pro Val Pro Ser Arg Ser Glu Ser Phe Ser Asn Gly Asn Ser
595 600 605

Glu Ser Val His Pro Ala Leu Gln Arg Pro Ala Glu Pro Gln Val Pro
610 615 620

Val Arg Thr Thr Ser Arg Ser Pro Val Leu Ser Arg Arg Asp Ser Pro

625	630	635	640
Leu Gln Gly Ser Gly Gln Gln Asn Ser Gln Ala Gly Gln Arg Asn Ser			
645	650	655	
Thr Ser Ile Glu Pro Arg Leu Leu Trp Glu Arg Val Glu Lys Leu Val			
660	665	670	
Pro Arg Pro Gly Ser Gly Ser Ser Gly Ser Ser Asn Ser Gly Ser			
675	680	685	
Gln Pro Gly Ser His Pro Gly Ser Gln Ser Gly Ser Gly -Glu Arg Phe			
690	695	700	
Arg Val Arg Ser Ser Lys Ser Glu Gly Ser Pro Ser Gln Arg Leu			
705	710	715	720
Glu Asn Ala Val Lys Lys Pro Glu Asp Lys Lys Glu Val Phe Arg Pro			
725	730	735	
Leu Lys Pro Ala Asp Leu Thr Ala Leu Ala Lys Glu Leu Arg Ala Val			
740	745	750	
Glu Asp Val Arg Pro Pro His Lys Val Thr Asp Tyr Ser Ser Ser			
755	760	765	
Glu Glu Ser Gly Thr Thr Asp Glu Glu Asp Asp Asp Val Glu Gln Glu			
770	775	780	
Gly Ala Asp Glu Ser Thr Ser Gly Pro Glu Asp Thr Arg Ala Ala Ser			
785	790	795	800
Ser Leu Asn Leu Ser Asn Gly Glu Thr Glu Ser Val Lys Thr Met Ile			
805	810	815	
Val His Asp Asp Val Glu Ser Glu Pro Ala Met Thr Pro Ser Lys Glu			
820	825	830	
Gly Thr Leu Ile Val Arg Arg Thr Gln Ser Ala Ser Ser Thr Leu Gln			
835	840	845	
Lys His Lys Ser Ser Ser Phe Thr Pro Phe Ile Asp Pro Arg Leu			
850	855	860	
Leu Gln Ile Ser Pro Ser Ser Gly Thr Thr Val Thr Ser Val Val Gly			
865	870	875	880
Phe Ser Cys Asp Gly Met Arg Pro Glu Ala Ile Arg Gln Asp Pro Thr			
885	890	895	
Arg Lys Gly Ser Val Val Asn Val Asn Pro Thr Asn Thr Arg Pro Gln			
900	905	910	
Ser Asp Thr Pro Glu Ile Arg Lys Tyr Lys Lys Arg Phe Asn Ser Glu			
915	920	925	

Ile Leu Cys Ala Ala Leu Trp Gly Val Asn Leu Leu Val Gly Thr Glu
930 935 940

Ser Gly Leu Met Leu Leu Asp Arg Ser Gly Gln Gly Lys Val Tyr Pro
945 950 955 960

Leu Ile Asn Arg Arg Arg Phe Gln Gln Met Asp Val Leu Glu Gly Leu
965 970 975

Asn Val Leu Val Thr Ile Ser Gly Lys Lys Asp Lys Leu Arg Val Tyr
980 985 990

Tyr Leu Ser Trp Leu Arg Asn Lys Ile Leu His Asn Asp Pro Glu Val
995 1000 1005

Glu Lys Lys Gln Gly Trp Thr Thr Val Gly Asp Leu Glu Gly Cys Val
1010 1015 1020

His Tyr Lys Val Val Lys Tyr Glu Arg Ile Lys Phe Leu Val Ile Ala
1025 1030 1035 1040

Leu Lys Ser Ser Val Glu Val Tyr Ala Trp Ala Pro Lys Pro Tyr His
1045 1050 1055

Lys Phe Met Ala Phe Lys Ser Phe Gly Glu Leu Val His Lys Pro Leu
1060 1065 1070

Leu Val Asp Leu Thr Val Glu Glu Gly Gln Arg Leu Lys Val Ile Tyr
1075 1080 1085

Gly Ser Cys Ala Gly Phe His Ala Val Asp Val Asp Ser Gly Ser Val
1090 1095 1100

Tyr Asp Ile Tyr Leu Pro Thr His Ile Gln Cys Ser Ile Lys Pro His
1105 1110 1115 1120

Ala Ile Ile Ile Leu Pro Asn Thr Asp Gly Met Glu Leu Leu Val Cys
1125 1130 1135

Tyr Glu Asp Glu Gly Val Tyr Val Asn Thr Tyr Gly Arg Ile Thr Lys
1140 1145 1150

Asp Val Val Leu Gln Trp Gly Glu Met Pro Thr Ser Val Ala Tyr Ile
1155 1160 1165

Arg Ser Asn Gln Thr Met Gly Trp Gly Glu Lys Ala Ile Glu Ile Arg
1170 1175 1180

Ser Val Glu Thr Gly His Leu Asp Gly Val Phe Met His Lys Arg Ala
1185 1190 1195 1200

Gln Arg Leu Lys Phe Leu Cys Glu Arg Asn Asp Lys Val Phe Phe Ala
1205 1210 1215

Ser Val Arg Ser Gly Gly Ser Ser Gln Val Tyr Phe Met Thr Leu Gly
1220 1225 1230

Arg Thr Ser Leu Leu Ser Trp
1235

<210> 14
<211> 1297
<212> PRT
<213> Mammalian (Human) ZC2

<400> 14

Ala Phe Gly Glu Val Tyr Glu Gly Arg His Val Lys Thr Gly Gln Leu
1 5 10 15

Ala Ala Ile Lys Val Met Asp Val Thr Gly Asp Glu Glu Glu Glu Ile
20 25 30

Lys Gln Glu Ile Asn Met Leu Lys Lys Tyr Ser His His Arg Asn Ile
35 40 45

Ala Thr Tyr Tyr Gly Ala Phe Ile Lys Lys Asn Pro Pro Gly Met Asp
50 55 60

Asp Gln Leu Trp Leu Val Met Glu Phe Cys Gly Ala Gly Ser Val Thr
65 70 75 80

Asp Leu Ile Lys Asn Thr Lys Gly Asn Thr Leu Lys Glu Glu Trp Ile
85 90 95

Ala Tyr Ile Cys Arg Glu Ile Leu Arg Gly Leu Ser His Leu His Gln
100 105 110

His Lys Val Ile His Arg Asp Ile Lys Gly Gln Asn Val Leu Leu Thr
115 120 125

Glu Asn Ala Glu Val Lys Leu Val Asp Phe Gly Val Ser Ala Gln Leu
130 135 140

Asp Arg Thr Val Gly Arg Arg Asn Thr Phe Ile Gly Thr Pro Tyr Trp
145 150 155 160

Met Ala Pro Glu Val Ile Ala Cys Asp Glu Asn Pro Asp Ala Thr Tyr
165 170 175

Asp Phe Lys Ser Asp Leu Trp Ser Leu Gly Ile Thr Ala Ile Glu Met
180 185 190

Ala Glu Gly Ala Pro Pro Leu Cys Asp Met His Pro Met Arg Ala Leu
195 200 205

Phe Leu Ile Pro Arg Asn Pro Ala Pro Arg Leu Lys Ser Lys Lys Trp

WO 99/53036

22

210	215	220
Ser Lys Lys Phe Gln Ser Phe Ile Glu Ser Cys Leu Val Lys Asn His		
230		235
Ser Gln Arg Pro Ala Thr Glu Gln Leu Met Lys His Pro Phe Ile Arg		240
245		250
Asp Gln Pro Asn Glu Arg Gln Val Arg Ile Gln Leu Lys Asp His Ile		255
260		265
Asp Arg Thr Lys Lys Lys Arg Gly Glu Lys Asp Glu Thr Glu Tyr Glu		270
275		280
Tyr Ser Gly Ser Glu Glu Glu Glu Asn Asp Ser Gly Glu Pro		285
290		295
Gln Gln Leu Glu Gln Gln Arg Glu Asn Glu Glu His Lys Arg Gln		300
325		340
Leu Arg Leu Gln Leu Ala Asn Lys Glu Arg Ser Glu Ala Leu Arg Arg		310
335		345
Arg Arg Leu Glu Glu Gln Gln Arg Arg Glu Lys Glu Leu Arg Lys Gln		320
355		360
Gln Glu Arg Glu Gln Arg Arg His Tyr Glu Glu Gln Lys Glu Gln Arg		330
375		380
Glu Glu Arg Arg Arg Ala Glu His Glu Gln Glu Tyr Lys Arg Lys Gln		350
405		410
Leu Glu Glu Gln Arg Gln Ala Glu Arg Leu Gln Arg Gln Leu Lys Gln		365
420		425
Glu Arg Asp Tyr Leu Val Ser Leu Gln His Gln Arg Gln Glu Gln Arg		380
435		440
Pro Val Glu Lys Lys Pro Leu Tyr His Tyr Lys Glu Gly Met Ser Pro		400
450		455
Ser Glu Lys Pro Ala Trp Ala Lys Glu Val Glu Glu Arg Ser Arg Leu		425
465		470
Asn Arg Gln Ser Ser Pro Ala Met Pro His Lys Val Ala Asn Arg Ile		445
485		490
Ser Asp Pro Asn Leu Pro Pro Arg Ser Glu Ser Phe Ser Ile Ser Gly		460
500		505
Ser Asp Pro Asn Leu Pro Pro Arg Ser Glu Ser Phe Ser Ile Ser Gly		480
505		510

Val Gln Pro Ala Arg Thr Pro Pro Met Leu Arg Pro Val Asp Pro Gln
515 520 525

Ile Pro His Leu Val Ala Val Lys Ser Gln Gly Pro Ala Leu Thr Ala
530 535 540

Ser Gln Ser Val His Glu Gln Pro Thr Lys Gly Leu Ser Gly Phe Gln
545 550 555 560

Glu Ala Leu Asn Val Thr Ser His Arg Val Glu Met Pro Arg Gln Asn
565 570 575

Ser Asp Pro Thr Ser Glu Asn Pro Pro Leu Pro Thr Arg Ile Glu Lys
580 585 590

Phe Asp Arg Ser Ser Trp Leu Arg Gln Glu Glu Asp Ile Pro Pro Lys
595 600 605

Val Pro Gln Arg Thr Thr Ser Ile Ser Pro Ala Leu Ala Arg Lys Asn
610 615 620

Ser Pro Gly Asn Gly Ser Ala Leu Gly Pro Arg Leu Gly Ser Gln Pro
625 630 635 640

Ile Arg Ala Ser Asn Pro Asp Leu Arg Arg Thr Glu Pro Ile Leu Glu
645 650 655

Ser Pro Leu Gln Arg Thr Ser Ser Gly Ser Ser Ser Ser Ser Thr
660 665 670

Pro Ser Ser Gln Pro Ser Ser Gln Gly Ser Gln Pro Gly Ser Gln
675 680 685

Ala Gly Ser Ser Glu Arg Thr Arg Val Arg Ala Asn Ser Lys Ser Glu
690 695 700

Gly Ser Pro Val Leu Pro His Glu Pro Ala Lys Val Lys Pro Glu Glu
705 710 715 720

Ser Arg Asp Ile Thr Arg Pro Ser Arg Pro Ala Ser Tyr Lys Lys Ala
725 730 735

Ile Asp Glu Asp Leu Thr Ala Leu Ala Lys Glu Leu Arg Glu Leu Arg
740 745 750

Ile Glu Glu Thr Asn Arg Pro Met Lys Lys Val Thr Asp Tyr Ser Ser
755 760 765

Ser Ser Glu Glu Ser Glu Ser Ser Glu Glu Glu Glu Asp Gly Glu
770 775 780

Ser Glu Thr His Asp Gly Thr Val Ala Val Ser Asp Ile Pro Arg Leu
785 790 795 800

Ile Pro Thr Gly Ala Pro Gly Ser Asn Glu Gln Tyr Asn Val Gly Met
805 810 815

Val Gly Thr His Gly Leu Glu Thr Ser His Ala Asp Ser Phe Ser Gly
820 825 830

Ser Ile Ser Arg Glu Gly Thr Leu Met Ile Arg Glu Thr Ser Gly Glu
835 840 845

Lys Lys Arg Ser Gly His Ser Asp Ser Asn Gly Phe Ala Gly His Ile
850 855 860

Asn Leu Pro Asp Leu Val Gln Gln Ser His Ser Pro Ala Gly Thr Pro
865 870 875 880

Thr Glu Gly Leu Gly Arg Val Ser Thr His Ser Gln Glu Met Asp Ser
885 890 895

Gly Thr Glu Tyr Gly Met Gly Ser Ser Thr Lys Ala Ser Phe Thr Pro
900 905 910

Phe Val Asp Pro Arg Val Tyr Gln Thr Ser Pro Thr Asp Glu Asp Glu
915 920 925

Glu Asp Glu Glu Ser Ser Ala Ala Ala Leu Phe Thr Gly Glu Leu Leu
930 935 940

Arg Gln Glu Gln Ala Lys Leu Asn Glu Ala Arg Lys Ile Ser Val Val
945 950 955 960

Asn Val Asn Pro Thr Asn Ile Arg Pro His Ser Asp Thr Pro Glu Ile
965 970 975

Arg Lys Tyr Lys Lys Arg Phe Asn Ser Glu Ile Leu Cys Ala Ala Leu
980 985 990

Trp Gly Val Asn Leu Leu Val Gly Thr Glu Asn Gly Leu Met Leu Leu
995 1000 1005

Asp Arg Ser Gly Gln Gly Lys Val Tyr Asn Leu Ile Asn Arg Arg Arg
1010 1015 1020

Phe Gln Gln Met Asp Val Leu Glu Gly Leu Asn Val Leu Val Thr Ile
1025 1030 1035 1040

Ser Gly Lys Lys Asn Lys Leu Arg Val Tyr Tyr Leu Ser Trp Leu Arg
1045 1050 1055

Asn Arg Ile Leu His Asn Asp Pro Glu Val Glu Lys Lys Gln Gly Trp
1060 1065 1070

Ile Thr Val Gly Asp Leu Glu Gly Cys Ile His Tyr Lys Val Val Lys
1075 1080 1085

Tyr Glu Arg Ile Lys Phe Leu Val Ile Ala Leu Lys Asn Ala Val Glu

1090

1095

1100

Ile Tyr Ala Trp Ala Pro Lys Pro Tyr His Lys Phe Met Ala Phe Lys
 1105 1110 1115 1120

Ser Phe Ala Asp Leu Gln His Lys Pro Leu Leu Val Asp Leu Thr Val
 1125 1130 1135

Glu Glu Gly Gln Arg Leu Lys Val Ile Phe Gly Ser His Thr Gly Phe
 1140 1145 1150

His Val Ile Asp Val Asp Ser Gly Asn Ser Tyr Asp Ile Tyr Thr Pro
 1155 1160 1165

Ser His Ile Gln Gly Asn Ile Thr Pro His Ala Ile Val Ile Leu Pro
 1170 1175 1180

Lys Thr Asp Gly Met Glu Met Leu Val Cys Tyr Glu Asp Glu Gly Val
 1185 1190 1195 1200

Tyr Val Asn Thr Tyr Gly Arg Ile Thr Lys Asp Val Val Leu Gln Trp
 1205 1210 1215

Gly Glu Met Pro Thr Ser Val Ala Tyr Ile His Ser Asn Gln Ile Met
 1220 1225 1230

Gly Trp Gly Glu Lys Ala Ile Glu Ile Arg Ser Val Glu Thr Gly His
 1235 1240 1245

Leu Asp Gly Val Phe Met His Lys Arg Ala Gln Arg Leu Lys Phe Leu
 1250 1255 1260

Cys Glu Arg Asn Asp Lys Val Phe Phe Ala Ser Val Arg Ser Gly Gly
 1265 1270 1275 1280

Ser Ser Gln Val Phe Phe Met Thr Leu Asn Arg Asn Ser Met Met Asn
 1285 1290 1295

Trp

<210> 15
 <211> 1326
 <212> PRT
 <213> Mammalian (Human) ZC3

<400> 15

Ala Phe Gly Glu Val Tyr Glu Gly Arg His Val Lys Thr Gly Gln Leu
 1 5 10 15

Ala Ala Ile Lys Val Met Asp Val Thr Glu Asp Glu Glu Glu Ile
 20 25 30

Lys Gln Glu Ile Asn Met Leu Lys Lys Tyr Ser His His Arg Asn Ile
 35 40 45
 Ala Thr Tyr Tyr Gly Ala Phe Ile Lys Lys Ser Pro Pro Gly Asn Asp
 50 55 60
 Asp Gln Leu Trp Leu Val Met Glu Phe Cys Gly Ala Gly Ser Val Thr
 65 70 75 80
 Asp Leu Val Lys Asn Thr Lys Gly Asn Ala Leu Lys Glu Asp Cys Ile
 85 90 95
 Ala Tyr Ile Cys Arg Glu Ile Leu Arg Gly Leu Ala His Leu His Ala
 100 105 110
 His Lys Val Ile His Arg Asp Ile Lys Gly Gln Asn Val Leu Leu Thr
 115 120 125
 Glu Asn Ala Glu Val Lys Leu Val Asp Phe Gly Val Ser Ala Gln Leu
 130 135 140
 Asp Arg Thr Val Gly Arg Arg Asn Thr Phe Ile Gly Thr Pro Tyr Trp
 145 150 155 160
 Met Ala Pro Glu Val Ile Ala Cys Asp Glu Asn Pro Asp Ala Thr Tyr
 165 170 175
 Asp Tyr Arg Ser Asp Ile Trp Ser Leu Gly Ile Thr Ala Ile Glu Met
 180 185 190
 Ala Glu Gly Ala Pro Pro Leu Cys Asp Met His Pro Met Arg Ala Leu
 195 200 205
 Phe Leu Ile Pro Arg Asn Pro Pro Pro Arg Leu Lys Ser Lys Lys Trp
 210 215 220
 Ser Lys Lys Phe Ile Asp Phe Ile Asp Thr Cys Leu Ile Lys Thr Tyr
 225 230 235 240
 Leu Ser Arg Pro Pro Thr Glu Gln Leu Leu Lys Phe Pro Phe Ile Arg
 245 250 255
 Asp Gln Pro Thr Glu Arg Gln Val Arg Ile Gln Leu Lys Asp His Ile
 260 265 270
 Asp Arg Ser Arg Lys Lys Arg Gly Glu Lys Glu Glu Thr Glu Tyr Glu
 275 280 285
 Tyr Ser Gly Ser Glu Glu Glu Asp Asp Ser His Gly Glu Glu Gly Glu
 290 295 300
 Pro Ser Ser Ile Met Asn Val Pro Gly Glu Ser Thr Leu Arg Arg Glu
 305 310 315 320
 Phe Leu Arg Leu Gln Gln Glu Asn Lys Ser Asn Ser Glu Ala Leu Lys
 325 330 335

Gln Gln Gln Gln Leu Gln Gln Gln Gln Arg Asp Pro Glu Ala His
340 345 350

Ile Lys His Leu Leu His Gln Arg Gln Arg Arg Ile Glu Glu Gln Lys
355 360 365

Glu Glu Arg Arg Val Glu Glu Gln Gln Arg Arg Glu Arg Glu Gln
370 375 380

Arg Lys Leu Gln Glu Lys Glu Gln Gln Arg Arg Leu Glu Asp Met Gln
385 390 395 400

Ala Leu Arg Arg Glu Glu Glu Arg Arg Gln Ala Glu Arg Glu Gln Glu
405 410 415

Tyr Ile Arg His Arg Leu Glu Glu Glu Gln Arg Gln Leu Glu Ile Leu
420 425 430

Gln Gln Gln Leu Leu Gln Glu Gln Ala Leu Leu Leu Glu Tyr Lys Arg
435 440 445

Lys Gln Leu Glu Glu Gln Arg Gln Ser Glu Arg Leu Gln Arg Gln Leu
450 455 460

Gln Gln Glu His Ala Tyr Leu Lys Ser Leu Gln Gln Gln Gln Gln
465 470 475 480

Gln Gln Leu Gln Lys Gln Gln Gln Gln Leu Leu Pro Gly Asp Arg
485 490 495

Lys Pro Leu Tyr His Tyr Gly Arg Gly Met Asn Pro Ala Asp Lys Pro
500 505 510

Ala Trp Ala Arg Glu Val Glu Glu Arg Thr Arg Met Asn Lys Gln Gln
515 520 525

Asn Ser Pro Leu Ala Lys Ser Lys Pro Gly Ser Thr Gly Pro Glu Pro
530 535 540

Pro Ile Pro Gln Ala Ser Pro Gly Pro Pro Gly Pro Leu Ser Gln Thr
545 550 555 560

Pro Pro Met Gln Arg Pro Val Glu Pro Gln Glu Gly Pro His Lys Ser
565 570 575

Leu Val Ala His Arg Val Pro Leu Lys Pro Tyr Ala Ala Pro Val Pro
580 585 590

Arg Ser Gln Ser Leu Gln Asp Gln Pro Thr Arg Asn Leu Ala Ala Phe
595 600 605

Pro Ala Ser His Asp Pro Asp Pro Ala Ile Pro Ala Pro Thr Ala Thr
610 615 620

Pro Ser Ala Arg Gly Ala Val Ile Arg Gln Asn Ser Asp Pro Thr Ser
625 630 635 640

Glu Gly Pro Gly Pro Ser Pro Asn Pro Pro Ala Trp Val Arg Pro Asp
645 650 655

Asn Glu Ala Pro Pro Lys Val Pro Gln Arg Thr Ser Ser Ile Ala Thr
660 665 670

Ala Leu Asn Thr Ser Gly Ala Gly Gly Ser Arg Pro Ala Gln Ala Val
675 680 685

Arg Ala Arg Pro Arg Ser Asn Ser Ala Trp Gln Ile Tyr Leu Gln Arg
690 695 700

Arg Ala Glu Arg Gly Thr Pro Lys Pro Pro Gly Pro Pro Ala Gln Pro
705 710 715 720

Pro Gly Pro Pro Asn Ala Ser Ser Asn Pro Asp Leu Arg Arg Ser Asp
725 730 735

Pro Gly Trp Glu Arg Ser Asp Ser Val Leu Pro Ala Ser His Gly His
740 745 750

Leu Pro Gln Ala Gly Ser Leu Glu Arg Asn Arg Val Gly Val Ser Ser
755 760 765

Lys Pro Asp Ser Ser Pro Val Leu Ser Pro Gly Asn Lys Ala Lys Pro
770 775 780

Asp Asp His Arg Ser Arg Pro Gly Arg Pro Ala Asp Phe Val Leu Leu
785 790 795 800

Lys Glu Arg Thr Leu Asp Glu Ala Pro Arg Pro Pro Lys Lys Ala Met
805 810 815

Asp Tyr Ser Ser Ser Glu Glu Val Glu Ser Ser Glu Asp Asp Glu
820 825 830

Glu Glu Gly Glu Gly Pro Ala Glu Gly Ser Arg Asp Thr Pro Gly
835 840 845

Gly Arg Asp Gly Asp Thr Asp Ser Val Ser Thr Met Val Val His Asp
850 855 860

Val Glu Glu Ile Thr Gly Thr Gln Pro Pro Tyr Gly Gly Thr Met
865 870 875 880

Val Val Gln Arg Thr Pro Glu Glu Arg Asn Leu Leu His Ala Asp
885 890 895

Ser Asn Gly Tyr Thr Asn Leu Pro Asp Val Val Gln Pro Ser His Ser
900 905 910

Pro Thr Glu Asn Ser Lys Gly Gln Ser Pro Pro Ser Lys Asp Gly Ser

915

920

925

Gly Asp Tyr Gln Ser Arg Gly Leu Val Lys Ala Pro Gly Lys Ser Ser
 930 935 940

Phe Thr Met Phe Val Asp Leu Gly Ile Tyr Gln Pro Gly Gly Ser Gly
 945 950 955 960

Asp Ser Ile Pro Ile Thr Ala Leu Val Gly Gly Glu Gly Thr Arg Leu
 965 970 975

Asp Gln Leu Gln Tyr Asp Val Arg Lys Gly Ser Val Val Asn Val Asn
 980 985 990

Pro Thr Asn Thr Arg Ala His Ser Glu Thr Pro Glu Ile Arg Lys Tyr
 995 1000 1005

Lys Lys Arg Phe Asn Ser Glu Ile Leu Cys Ala Ala Leu Trp Gly Val
 1010 1015 1020

Asn Leu Leu Val Gly Thr Glu Asn Gly Leu Met Leu Leu Asp Arg Ser
 1025 1030 1035 1040

Gly Gln Gly Lys Val Tyr Gly Leu Ile Gly Arg Arg Arg Phe Gln Gln
 1045 1050 1055

Met Asp Val Leu Glu Gly Leu Asn Leu Leu Ile Thr Ile Ser Gly Lys
 1060 1065 1070

Arg Asn Lys Leu Arg Val Tyr Tyr Leu Ser Trp Leu Arg Asn Lys Ile
 1075 1080 1085

Leu His Asn Asp Pro Glu Val Glu Lys Lys Gln Gly Trp Thr Thr Val
 1090 1095 1100

Gly Asp Met Glu Gly Cys Gly His Tyr Arg Val Val Lys Tyr Glu Arg
 1105 1110 1115 1120

Ile Lys Phe Leu Val Ile Ala Leu Lys Ser Ser Val Glu Val Tyr Ala
 1125 1130 1135

Trp Ala Pro Lys Pro Tyr His Lys Phe Met Ala Phe Lys Ser Phe Ala
 1140 1145 1150

Asp Leu Pro His Arg Pro Leu Leu Val Asp Leu Thr Val Glu Glu Gly
 1155 1160 1165

Gln Arg Leu Lys Val Ile Tyr Gly Ser Ser Ala Gly Phe His Ala Val
 1170 1175 1180

Asp Val Asp Ser Gly Asn Ser Tyr Asp Ile Tyr Ile Pro Val His Ile
 1185 1190 1195 1200

Gln Ser Gln Ile Thr Pro His Ala Ile Ile Phe Leu Pro Asn Thr Asp

30

1205

1210

1215

Gly Met Glu Met Leu Leu Cys Tyr Glu Asp Glu Gly Val Tyr Val Asn
1220 1225 1230

Thr Tyr Gly Arg Ile Ile Lys Asp Val Val Leu Gln Trp Gly Glu Met
1235 1240 1245

Pro Thr Ser Val Ala Tyr Ile Cys Ser Asn Gln Ile Met Gly Trp Gly
1250 1255 1260

Glu Lys Ala Ile Glu Ile Arg Ser Val Glu Thr Gly His Leu Asp Gly
1265 1270 1275 1280

Val Phe Met His Lys Arg Ala Gln Arg Leu Lys Phe Leu Cys Glu Arg
1285 1290 1295

Asn Asp Lys Val Phe Phe Ala Ser Val Arg Ser Gly Gly Ser Ser Gln
1300 1305 1310

Val Tyr Phe Met Thr Leu Asn Arg Asn Arg Ile Met Asn Trp
1315 1320 1325

<210> 16

<211> 349

<212> PRT

<213> Mammalian (Human) ZC4

<400> 16

Asn Val Asn Pro Leu Tyr Val Ser Pro Ala Cys Lys Lys Pro Leu Ile
1 5 10 15

His Met Tyr Glu Lys Glu Phe Thr Ser Glu Ile Cys Cys Gly Ser Leu
20 25 30

Trp Gly Val Asn Leu Leu Gly Thr Arg Ser Asn Leu Tyr Leu Met
35 40 45

Asp Arg Ser Gly Lys Ala Asp Ile Thr Lys Leu Ile Arg Arg Arg Pro
50 55 60

Phe Arg Gln Ile Gln Val Leu Glu Pro Leu Asn Leu Ile Thr Ile
65 70 75 80

Ser Gly His Lys Asn Arg Leu Arg Val Tyr His Leu Thr Trp Leu Arg
85 90 95

Asn Lys Ile Leu Asn Asn Asp Pro Glu Ser Lys Arg Arg Gln Glu Glu
100 105 110

Met Leu Lys Thr Glu Glu Ala Cys Lys Ala Ile Asp Lys Leu Thr Gly
115 120 125

31

Cys Glu His Phe Ser Val Leu Gln His Glu Glu Thr Thr Tyr Ile Ala
 130 135 140

Ile Ala Leu Lys Ser Ser Ile His Leu Tyr Ala Trp Ala Pro Lys Ser
 145 150 155 160

Phe Asp Glu Ser Thr Ala Ile Lys Val Phe Pro Thr Leu Asp His Lys
 165 170 175

Pro Val Thr Val Asp Leu Ala Ile Gly Ser Glu Lys Arg Leu Lys Ile
 180 185 190

Phe Phe Ser Ser Ala Asp Gly Tyr His Leu Ile Asp Ala Glu Ser Glu
 195 200 205

Val Met Ser Asp Val Thr Leu Pro Lys Asn Pro Leu Glu Ile Ile Ile
 210 215 220

Pro Gln Asn Ile Ile Ile Leu Pro Asp Cys Leu Gly Ile Gly Met Met
 225 230 235 240

Leu Thr Phe Asn Ala Glu Ala Leu Ser Val Glu Ala Asn Glu Gln Leu
 245 250 255

Phe Lys Lys Ile Leu Glu Met Trp Lys Asp Ile Pro Ser Ser Ile Ala
 260 265 270

Phe Glu Cys Thr Gln Arg Thr Thr Gly Trp Gly Gln Lys Ala Ile Glu
 275 280 285

Val Arg Ser Leu Gln Ser Arg Val Leu Glu Ser Glu Leu Lys Arg Arg
 290 295 300

Ser Ile Lys Lys Leu Arg Phe Leu Cys Thr Arg Gly Asp Lys Leu Phe
 305 310 315 320

Phe Thr Ser Thr Leu Arg Asn His His Ser Arg Val Tyr Phe Met Thr
 325 330 335

Leu Gly Lys Leu Glu Glu Leu Gln Ser Asn Tyr Asp Val
 340 345

<210> 17

<211> 4023

<212> DNA

<213> Mammalian (Human) KHS2

<400> 17

ccggccatgaa cccccggcttc gatttgcggcc gcccggaaaccc gcaggaggac ttcgagctga	60
ttcagcgcat cggcagcgcc acctacggcg acgtctacaa ggcacggaat gttaacactg	120
gtgaatttagc agcaattaaa gtaataaaaat tggaccagg agaagacttt gcagttgtgc	180
agcaagaaaat tattatgttgaa aacacccaaa tatttgtgct tattttggaa	240

gctatctcag	gcgagataag	cttggattt	gcatggagtt	ttgtggaggt	ggttctttac	300
aggatattta	tcaacgtact	ggacctctgt	cagaactgca	aattgcata	tttagcagag	360
aaacactgca	gggattata	tatcttcaca	gtaaaggaaa	aatgcacaga	gatataaagg	420
gagctaaca	tctattaacg	gataatggc	atgtgaaattt	ggctgatttt	ggagtatctg	480
cacagataac	agctacaattt	gccaaacgga	agtctttcat	tggcacacca	tattggatgg	540
ctccagaagt	tgcagctgtt	gagagaagg	ggggttacaa	tcaactctgt	gatctctggg	600
cagtggaaat	cactgccata	gaactgcag	agcttcagcc	tcctatgttt	gacttacacc	660
caatgagagc	attatttcta	atgacaaaaaa	gcaattttca	gcctcctaaa	ctaaaggata	720
aatgaaatg	gtcaaata	gttcatca	ttgtgaaaat	ggcacttacc	aaaaatccga	780
aaaaaagacc	tactgctgaa	aaattattac	agcatcctt	tgtaacacaa	catttgacac	840
ggtcttggc	aatcgagctg	ttggataaaag	taaataatcc	agatcattcc	acttaccatg	900
atttcgatga	tgatgatcct	gagcctcttg	ttgctgtacc	acatagaatt	cactcaacaa	960
gtagaaacgt	gagagaagaa	aaaacacgct	cagagataac	cttggccaa	gtgaaatttg	1020
atccaccctt	aagaaaggag	acagaaccac	atcatgaact	tcccgcacgt	gatggtttt	1080
tggacagttc	agaagaata	tactacactg	caagatctaa	tctggatctg	caactggaat	1140
atggacaagg	acaccaaggt	gtttaacttt	tagtgcaaa	caagagtctt	ctcaagtctg	1200
ttgaagaaga	attgcatacg	cgaggacacg	tcgcacattt	agaagatgt	gaaggagatg	1260
atgatgaatc	taaacactca	actctgaaag	caaaaattcc	accccttgg	ccacccaagc	1320
ctaagtcata	cttcataccca	caggaaatgc	attctactga	ggatgaaaat	caaggaacaa	1380
tcaagagatg	tcccatgtca	gggagccag	caaagccatc	ccaagttcca	cctagaccac	1440
cacctccag	attacccccca	cacaaacctg	ttgccttagg	aatggaaatg	agctccctcc	1500
agttaaatgg	tgaacgagat	ggctcattat	gtcaacaaca	gaatgaacat	agaggcaca	1560
accttcaag	aaaagaaaag	aaagatgtac	caaagcctat	tagtaatgg	cttcctccaa	1620
cacctaaatg	gcatatgggt	gcatgtttt	caaagtttt	taatgggtgt	cccttggaaa	1680
ttcactgtgc	atcatcatgg	ataaaacccag	atacaagaga	tcaacttgc	atattttgtg	1740
ccgaagaagg	gatttatacc	ctcaatctt	atgaacttca	tgaacatca	atggAACAGC	1800
tattccctcg	aagggttaca	ttgttgtatg	taatgaacaa	ttgttgtcta	tcaatatctg	1860
gtaaagcttc	tcaagtttat	tccctataatt	taccaggct	ttttgattat	gcaagacaaa	1920
tgcaaaagtt	acctgttgc	attccagcac	acaaactccc	tgacagaata	ctgccaagga	1980
aattttctgt	atcagcaaaa	atccctgaaa	ccaaatggtg	ccagaagtgt	tgtgtgtaa	2040
gaaatccta	cacggccat	aaataccat	gtggagca	tcaacttgc	attgttctat	2100
tagaatgggt	tgaaccaatg	cagaaattta	ttgttaattaa	gcacatagat	tttcctatac	2160
catgtccact	tagaatgttt	gaaatgttgg	tagttccat	acaggagtt	cctttatgtt	2220
gtgttgtgt	cagtagaggt	agagactca	accaagttgt	tcgatttgag	acggtaatc	2280
caaattctac	ctcttcatgg	tttacagaat	cagatacccc	acagacaaaat	gttactcatg	2340
taacccaact	ggagagagat	accatccttgc	tatgtttgg	ctgttgtata	aaaatagtaa	2400
atctccaagg	aagattaaaa	tctacggaga	aattgtcata	agaactcacc	tttgatttcc	2460
agattgaatc	aatagtgtc	ctacaagaca	gtgtgttagc	tttctggaaa	catggatgc	2520
aaggtagaag	ttttagatct	aatgaggtaa	cacaagaaat	ttcagatagc	acaagaattt	2580
tcaggctgct	tggatctgac	agggtcggtt	ttttggaaag	taggcaact	gataacccca	2640
cagcaaatag	caatttgcac	atccctggcg	gtcatgaaaa	cagttactga	gaattgtgt	2700
gctttgacag	ttaactctag	aaagaaagaa	cactaccact	gcaacattaa	ttggatgttg	2760
aagctgtaca	aaagctgcag	taacctgtct	tcaacttgc	tgttattat	tgtggcatga	2820
gataagatgg	ggaaaatttt	gttttaagt	gtatggat	atttagcata	ttgaaccaca	2880
caagtgccta	attcattgtt	atgtaatctt	tgtacatata	ggcagtattt	tttctgtgaa	2940
acttcataatt	gctgaagaca	tacactaaga	atttatgtat	ataatgtact	tttatgagat	3000
gtacaagtaa	gtgtcttata	tgtacagatg	taaatgtga	tggaggctt	tttgcata	3060
atatttttaag	aattctttag	tatattttgc	ggtgtggct	tattacaaaa	ttggatgttg	3120
gcaatgaaac	aatacatttta	acactattgt	atttttata	tatgttaattt	agtaatatga	3180
atataaaatct	tgtactttt	aaaattgtaa	tggaggctt	aatcatttttta	taatctttt	3240
aatttttaatg	caagtacact	gggtttata	tttgcacaaa	gtattgtat	gtgatgtatt	3300
aagtccacaa	agtaagctgt	gacattgtct	ataagcattt	ggctccacaa	atgtatttgg	3360
attttttct	atgtgaagca	aaccaattat	aattaaccac	atgttgtagt	aactggctt	3420
tttatattta	agcagaatcc	tgttaagattt	tttgttttg	cttaaaaaca	atacccttga	3480
acatttttga	atcacagaat	agcggtacca	tgtatagaata	ctgcaattgt	ggtcagaattt	3540
acagatgtca	caaagaat	attagcatta	ttaaagagtc	ctcaactaaac	atttcatatg	3600

atcacactga	agaactgtaa	cattccatag	agtgaagtgg	ttcaaatttc	tcttggatt	3660
tttacttttg	ttggccttat	tttatgatcc	tttcatatt	tctttgact	tagaggatta	3720
atacatggcc	aaaataattt	agttactacc	tcatacaaaac	aatataatgg	ttactacaca	3780
tcacaggaac	ttagtttgg	tttaagtcat	tttgattgc	tttttccaa	tggaatatgt	3840
atataccagg	tttagcaaa	atgcacactt	ttggctctt	ttggtatatg	ttctttatat	3900
ttaatgtga	gtatatacac	taagaacaaa	ctaaattgtg	atttatgatc	ttcatttatt	3960
ttaatgataa	tggtttaaa	atatgttcct	gattgtacat	attgtaaaat	aaacatgttt	4020
ttt						4023

<210> 18
<211> 894
<212> PRT
<213> Mammalian (Human) KHS2

<400> 18

Met Asn Pro Gly Phe Asp Leu Ser Arg Arg Asn Pro Gln Glu Asp Phe
1 5 10 15

Glu Leu Ile Gln Arg Ile Gly Ser Gly Thr Tyr Gly Asp Val Tyr Lys
20 25 30

Ala Arg Asn Val Asn Thr Gly Glu Leu Ala Ala Ile Lys Val Ile Lys
35 40 45

Leu Glu Pro Gly Glu Asp Phe Ala Val Val Gln Gln Glu Ile Ile Met
50 55 60

Met Lys Asp Cys Lys His Pro Asn Ile Val Ala Tyr Phe Gly Ser Tyr
65 70 75 80

Leu Arg Arg Asp Lys Leu Trp Ile Cys Met Glu Phe Cys Gly Gly
85 90 95

Ser Leu Gln Asp Ile Tyr His Val Thr Gly Pro Leu Ser Glu Leu Gln
100 105 110

Ile Ala Tyr Val Ser Arg Glu Thr Leu Gln Gly Leu Tyr Tyr Leu His
115 120 125

Ser Lys Gly Lys Met His Arg Asp Ile Lys Gly Ala Asn Ile Leu Leu
130 135 140

Thr Asp Asn Gly His Val Lys Leu Ala Asp Phe Gly Val Ser Ala Gln
145 150 155 160

Ile Thr Ala Thr Ile Ala Lys Arg Lys Ser Phe Ile Gly Thr Pro Tyr
165 170 175

Trp Met Ala Pro Glu Val Ala Ala Val Glu Arg Lys Gly Gly Tyr Asn
180 185 190

Gln Leu Cys Asp Leu Trp Ala Val Gly Ile Thr Ala Ile Glu Leu Ala
195 200 205

Glu Leu Gln Pro Pro Met Phe Asp Leu His Pro Met Arg Ala Leu Phe
 210 215 220
 Leu Met Thr Lys Ser Asn Phe Gln Pro Pro Lys Leu Lys Asp Lys Met
 225 230 235 240
 Lys Trp Ser Asn Ser Phe His His Phe Val Lys Met Ala Leu Thr Lys
 245 250 255
 Asn Pro Lys Lys Arg Pro Thr Ala Glu Lys Leu Leu Gln His Pro Phe
 260 265 270
 Val Thr Gln His Leu Thr Arg Ser Leu Ala Ile Glu Leu Leu Asp Lys
 275 280 285
 Val Asn Asn Pro Asp His Ser Thr Tyr His Asp Phe Asp Asp Asp Asp
 290 295 300
 Pro Glu Pro Leu Val Ala Val Pro His Arg Ile His Ser Thr Ser Arg
 305 310 315 320
 Asn Val Arg Glu Glu Lys Thr Arg Ser Glu Ile Thr Phe Gly Gln Val
 325 330 335
 Lys Phe Asp Pro Pro Leu Arg Lys Glu Thr Glu Pro His His Glu Leu
 340 345 350
 Pro Asp Ser Asp Gly Phe Leu Asp Ser Ser Glu Glu Ile Tyr Tyr Thr
 355 360 365
 Ala Arg Ser Asn Leu Asp Leu Gln Leu Glu Tyr Gly Gln Gly His Gln
 370 375 380
 Gly Gly Tyr Phe Leu Gly Ala Asn Lys Ser Leu Leu Lys Ser Val Glu
 385 390 395 400
 Glu Glu Leu His Gln Arg Gly His Val Ala His Leu Glu Asp Asp Glu
 405 410 415
 Gly Asp Asp Asp Glu Ser Lys His Ser Thr Leu Lys Ala Lys Ile Pro
 420 425 430
 Pro Pro Leu Pro Pro Lys Pro Lys Ser Ile Phe Ile Pro Gln Glu Met
 435 440 445
 His Ser Thr Glu Asp Glu Asn Gln Gly Thr Ile Lys Arg Cys Pro Met
 450 455 460
 Ser Gly Ser Pro Ala Lys Pro Ser Gln Val Pro Pro Arg Pro Pro Pro
 465 470 475 480
 Pro Arg Leu Pro Pro His Lys Pro Val Ala Leu Gly Asn Gly Met Ser
 485 490 495
 Ser Phe Gln Leu Asn Gly Glu Arg Asp Gly Ser Leu Cys Gln Gln

35

500

505

510

Asn Glu His Arg Gly Thr Asn Leu Ser Arg Lys Glu Lys Lys Asp Val
 515 520 525

Pro Lys Pro Ile Ser Asn Gly Leu Pro Pro Thr Pro Lys Val His Met
 530 535 540

Gly Ala Cys Phe Ser Lys Val Phe Asn Gly Cys Pro Leu Lys Ile His
 545 550 555 560

Cys Ala Ser Ser Trp Ile Asn Pro Asp Thr Arg Asp Gln Tyr Leu Ile
 565 570 575

Phe Gly Ala Glu Glu Gly Ile Tyr Thr Leu Asn Leu Asn Glu Leu His
 580 585 590

Glu Thr Ser Met Glu Gln Leu Phe Pro Arg Arg Cys Thr Trp Leu Tyr
 595 600 605

Val Met Asn Asn Cys Leu Leu Ser Ile Ser Gly Lys Ala Ser Gln Leu
 610 615 620

Tyr Ser His Asn Leu Pro Gly Leu Phe Asp Tyr Ala Arg Gln Met Gln
 625 630 635 640

Lys Leu Pro Val Ala Ile Pro Ala His Lys Leu Pro Asp Arg Ile Leu
 645 650 655

Pro Arg Lys Phe Ser Val Ser Ala Lys Ile Pro Glu Thr Lys Trp Cys
 660 665 670

Gln Lys Cys Cys Val Val Arg Asn Pro Tyr Thr Gly His Lys Tyr Leu
 675 680 685

Cys Gly Ala Leu Gln Thr Ser Ile Val Leu Leu Glu Trp Val Glu Pro
 690 695 700

Met Gln Lys Phe Met Leu Ile Lys His Ile Asp Phe Pro Ile Pro Cys
 705 710 715 720

Pro Leu Arg Met Phe Glu Met Leu Val Val Pro Glu Gln Glu Tyr Pro
 725 730 735

Leu Val Cys Val Gly Val Ser Arg Gly Arg Asp Phe Asn Gln Val Val
 740 745 750

Arg Phe Glu Thr Val Asn Pro Asn Ser Thr Ser Ser Trp Phe Thr Glu
 755 760 765

Ser Asp Thr Pro Gln Thr Asn Val Thr His Val Thr Gln Leu Glu Arg
 770 775 780

Asp Thr Ile Leu Val Cys Leu Asp Cys Cys Ile Lys Ile Val Asn Leu
 785 790 795 800

Gln Gly Arg Leu Lys Ser Ser Arg Lys Leu Ser Ser Glu Leu Thr Phe
 805 810 815

Asp Phe Gln Ile Glu Ser Ile Val Cys Leu Gln Asp Ser Val Leu Ala
 820 825 830

Phe Trp Lys His Gly Met Gln Gly Arg Ser Phe Arg Ser Asn Glu Val
 835 840 845

Thr Gln Glu Ile Ser Asp Ser Thr Arg Ile Phe Arg Leu Leu Gly Ser
 850 855 860

Asp Arg Val Val Val Leu Glu Ser Arg Pro Thr Asp Asn Pro Thr Ala
 865 870 875 880

Asn Ser Asn Leu Tyr Ile Leu Ala Gly His Glu Asn Ser Tyr
 885 890

<210> 19

<211> 4196

<212> DNA

<213> Mammalian (Human) SULU1

<400> 19

gggagggtcc	ttgtggcgcc	gggcggcggg	gtcctgcgtg	gagagtggga	cgcaacgccc	60
agaccgcgag	cagaggctgc	gcacagccgg	atccggcaact	cagcgaccgg	acccaaggat	120
ccgcccggga	acaaggccaca	ggagagcgac	tcaggaacaa	gtgtgggaga	ggaagcggcg	180
gcggcgccgc	cgggccccggg	ggtgggtgaca	gcaggtctga	ggttgcata	taaatacaaa	240
ggactgaagt	tataaaagaa	aaaagagaag	tttgctgcta	aatgaatct	gagcaatatg	300
gaatattttg	tgcccacacac	aaaaaggat	tgaagattta	ccccccaaaa	aaaatttgtca	360
atgagaaaata	aagctaactg	atataaaaaa	gcagagcctg	ctctactggc	catcatgcgt	420
aaaggggtgc	tgaaggacc	agagattgac	gatctattct	acaaagatga	tcctgaggaa	480
cttttattt	gtttgcata	aattggacat	ggaagttttg	gaggagtttta	ttttgcata	540
aatgctcaca	ccaatgagg	gttggcaatt	aagaagatgt	cctatagtgg	gaagcagacc	600
catgagaaat	ggcaagat	tcttaaggaa	gttaaatttt	tacgacaatt	gaagcatcct	660
aatactattt	agtacaaagg	ctgttactt	aaagaacaca	ctgttggtt	ggtgatggaa	720
tattgcctt	gtcagcctc	tgatttatta	gaagttcata	aaaaaccact	tcaggaagt	780
gagatcgctg	ccattactca	tggagcctt	catggactag	cctaccttaca	ttctcatgca	840
ttgattcata	gggatattaa	agcagggaaat	attcttctaa	cagagccagg	tcaggtaaaa	900
ctagctgatt	ttggatctgc	ttcaatggct	tctcctggca	actcttttgt	gggcacac	960
tactggatgg	ctccagaggt	gatcttagct	atggatgaag	gacaatgt	tgaaaagtt	1020
gatattttgt	caacttggcat	cacttgttatt	gaattggcg	aacggaaagcc	gcccttttc	1080
aacatgaatg	caatgagtgc	tttatatcac	attgcccaga	atgactcccc	aacgttacag	1140
tctaataat	ggacagactc	ttttaggaga	tttggatt	actgcgttgca	gaaaatacct	1200
caggaaaggc	caacatcagc	agaactatta	aggcatgact	ttttcgacq	agaccggcca	1260
ctacgtgtcc	tcattgac	catacagagg	acaaaagatg	cagtgcgtga	gctagataac	1320
ctacagtacc	aaaaatgaa	aaaaatactt	ttccaagaga	cacggaatgg	acccttgaat	1380
gagtcacagg	aggatgagga	agacagt	catggaacca	gcctgaacag	ggaaatggac	1440
agcctggc	gcaaccattc	cattccaage	atgtccgtga	gcacaggcag	ccagagcagc	1500
agtgtgaaca	gcatgcagga	agtcatggac	gagagcagtt	ccgaacttgt	catgatgcac	1560
gatgacgaaa	gcacaatcaa	ttccagctcc	tccgtcg	ataagaaaaga	tcatgtattc	1620

<210> 20

<211> 3824

<212> DNA

<213> Mammalian (Human) sun

<400> 20

tattgaattg gcggAACcgga agccttccttt atttaaatatg aatgcAATGA gtgcCttata
tcacatAGcc caaaatgaat cccctacact acagtctaAt gaatggtctg attattttcg 60
caactttgtA gattcttgcC tccagaaaat ccctcaagat cgacctacat cagaggaact 120
180

tttaaagcac atatttggc ttcgggagcg ccctgaaacc gtgttaatag atctcattca
 gaggacaaag gatgcgatcaa gagagctgga caatctgcag tatcgaaaga tgaagaaaact
 cctttccag gaggcacata atggaccagc agtagaaagca caggaagaag aagaggaaca
 agatcatgtt gttggccgga caggaacagt taatagtgtt ggaagtaatc aatccattcc
 cagcatgtcc atcagtgcac gcagccaaag cagtagtgc aacagtctt cagatgtctc
 agatgacaag agtggactg acatgtggaa gggagaccac acagtgtgt ctaacagttc
 tggttatccat taaaaccag aggaagaaaa ttacagagaa gaggagatc ctagaacaag
 agcatcgat ccacaatctc caccctcaatc actcgatcg aggcaatgc aagaacatga
 gcaggactct gagcttagag aacaatgtc tggctataag cgaatgtg gcaatgtc
 aaagcaactg atgactctgg aaaacaagct aaaggcttag atggatgaac atcgctc
 attagacaaa gatctgaaa ctacgttcaatc gcttggatgg agaaacttat
 caagaaacac caggctgcca tgagaaaga gcttggatgg gcaatgtg aagagaaaa
 atttcagcaa catattcagg cccaacagaa gaaagaactg aatagtttc tcgatccc
 gaaaagagag tataaacttc gaaaagagca gcttggatgg gagcttggatgg aagagaaaa
 tacccttcaaa aaagaaaaac aggatgttcc ttcaaaagc aagggatc aaaaaccag
 ccaagcagaa gaagaagcta acittcttc acgttcaaga aataccatc agctggatc
 ccgtcgcttc aagagaagaa tggttacttgg gcttggatgg accttgc
 ggaggagttt aacaaaagac agactcagaa ggacttagat catggatgg
 gcatgaatct atgcaagaac tggatgttcc gcttggatgg acatccatc
 tgatgttgc agattacagc atcacaactgat catggatgg
 aagagaacga gaactaagac gaaagcatgtt gttcaggat
 gaagtctaaa gaactccaaa taaaagca gttcaggat
 acagtacaaa gcattaagaa atcacaactgat
 tggatgttgc cggctcaagg aggaacagac
 tgatcacagc attaatgaaa tgctctccac
 agcagagtgc caggtttga agatgcagc
 tcagagcaaa atcaagatgc aagctgaggc
 acagagggtc tccctccgga gggcactt
 tttcagaaat gagcgcacag aacgataac
 tgaagcttt gactctgaaa gcatggact
 cccccctgag gcattcagcc
 tccaaatgca ggtggacccc
 tcgaggtgc agtatgggag
 gtcacacatg tttatacat
 tggaagctga gtgcataatgg
 tactaattgg gatgtcatag
 aaaagtggag ttgatattaa
 acacatatat tatcatgtt
 gcaaaaatag aacgtgacaa
 aagttctga aaagatctaa
 acagcaggca gcatatgtt
 gtgttattat ttgttattaa
 aagcagaat ccatgaacac
 ccatgaacac taaaaggactt
 tttttctg aatgcgtcat
 ttgttattat ataagatgg
 acaaaccat gggcttcaa
 tatggtggtt atgtcccgca
 ttgttatac ttgtatct
 tactttact ttgttgc
 ggtgttgc
 cataattgtt atctccctt
 38

atttaaatgt tttagttgcca cagtgaacca ggagtcactg agccaatgac tttaccagct	3600
gctgactaat ctccatcacc actgttagatt ttgctgcatt tgccaggctt ctatTTTAA	3660
ttgctgtttt cgttgcgtca gtactttaca aacttctagt tcgttgagac tttagtgacca	3720
tttggcatca agttaaacatc acacaatagg aaacaccact tccacaagtc tcaaggctca	3780
gtgctaaagt actactgaaa aggaactagg aagtttgcc aatt	3824

<210> 21
<211> 2249
<212> DNA
<213> Mammalian (Murine) SULU3

<400> 21

gcaggatgcc atcaactaac agagcaggca gtctaaagga ccctgaaatt gcagagctct	60
tcttcaaaga agatccggaa aagctttca cagatctcag agaaaatcgcc catggggact	120
ttggaggcgt atatTTGCA cgagatgtgc gtactaatga agtggggcc atcaagaaaa	180
tgtcttatAG tggaaagcag tctactgaga aatggcagga tattattaag gaagtcagt	240
ttctacAAAG aataaaaacat cccaacagta tagaatacaa aggctgctat ttacgtgaac	300
acacagcatg gcttgtaatg gaatattgtt taggatctgc ttccagattt tttagaagttc	360
ataaaaaAGCC attacaagaa gtggaaatag cagcaattac acatgggtct ctccaggAAC	420
tagtttACATT acatttcatt accatgatcc atagagatat caaaggcagga aatatCCTC	480
tgacagaACC aggccaagtgg aaacttgcgt accttggatc tgcttccatg gcttcccTG	540
ccaaTTCTT tggggaaaca ccatttggaa tggccccAGA agtaattttt gccatggatg	600
aaggacAGTA tggatggcaaa gttgatgtat ggtctcttgg aataacgtgt attgaattAG	660
ccgagaggAA gcctccTTA tttatATGA atgcaatgag tgccttatAT cacatAGCCC	720
aaaatGAATC ccctacACTA caatctaATA tgaatgatTC ttgcctccAG aaaatCCCTC	780
aagatCGCCC tacatcAGAG gaactttAA agcacatgtt tgTTCTTCA gaggcccTG	840
aaacAGTGTt aatAGATCTT attCAAAGGA caaaggatgc agtaagAGAG ctggacaATC	900
tgcagtATCG aaAGATGAAG aaactccTTT tccaggAGGC acataatGGG ccagcGGTAG	960
aagcacAGGA agaAGAGGG gagcaAGATC atggTGTGG ccgaACAGGA acAGTGAATA	1020
gtgttggAAAG caatcAGTCT atccCTAGTA tgcTATCAG tgccAGCAGT caaAGCAGCA	1080
gtgttaATAG tcttccAGAT gcatcAGATG acaAGAGTG gctAGACATG atggAGGGAG	1140
accatacAGT gatgtCTAAC agttcGTCA tccacttAAA acctgAGGGAG gaaaattACC	1200
aggaAGAAGG agatcCTAGA acaAGAGCAT cagACCCACA gtctcccccT caggGTCTC	1260
gtcacaAGTC acattATCGT aatAGAGAAC acTTTGCACAC catacGAACA gcatcACTGG	1320
ttacaAGACA gatgcaAGAA catgAGCAGG actCTGAACt tagAGAAACAG atgtCTGGTT	1380
ataAGCGGAT gaggcGACAG catcaAAAGC agctgatgac gctggAAAT aaactGAAGG	1440
cagAGATGGa cgaACATCGG ctcAGATTAG acaaAGATCT tgAAACTCAG cgtAAACAATT	1500
tcgCTGCAGA aatGGAGAAA ctTATTAAGA aacACCAAGC tgCTATGGAA aaAGAGGCTA	1560
aagtGATGGC caatGAGGAG aaaaaATTCC agcaACACAT tcAGGCTCAA cagaaaaAAAG	1620
aactGAATAG ctTTTGGAG tctcaAAAAA gagaATATAA actTCGCAAAG gagcAGCTTA	1680
aggAGGAGCT gaatGAAAAC cagAGCACAC ctAAAAAAGA aaAGCAGGAA tggCTTCAAA	1740
agcAGAAAGGA gaatATACAG cattttCAGG cagaAGAAAG agctaATCTT ctTCGACGTC	1800
aaaggcAGTA tctAGAGCTA gaatGTCGTc gcttcaAAAG aagaATGTTA ctTGGGCGAC	1860
ataactTGGA acaggACCTT gtcAGGGAGG agttAAACAA aaggcAGACT caAAAGGACT	1920
tggAACATGC aatGCTATTG cgACAGCATG aatCAATGCA agaACTGGAG tttcGCCATC	1980
tcaacACTAT tcAGAAGATG cgCTGTGAGT tgATCAGACT gcAGCATCAA actGAGCTCA	2040
ctaaccAGCT agAGTACAAT aAGAGAAGGG AACGGGAACt gaggcAAAGG catGTCAAGG	2100
aagttcGACA acaACCTAAG agtCTGAAGT ctaAAAGAACT ccaaATAAAA aAGCAGTTc	2160
aggataACCTG caaaATTCAA accAGACAGT acaaAGCATT aaggAAATCAC ctACTGGAGA	2220
ctacacAAAGGAGCAC aaAGCAATC	2249

<210> 22
<211> 898

40

<212> PRT

<213> Mammalian (Human) SULU1

<400> 22

Met Arg Lys Gly Val Leu Lys Asp Pro Glu Ile Asp Asp Leu Phe Tyr
1 5 10 15

Lys Asp Asp Pro Glu Glu Leu Phe Ile Gly Leu His Glu Ile Gly His
20 25 30

Gly Ser Phe Gly Ala Val Tyr Phe Ala Thr Asn Ala His Thr Asn Glu
35 40 45

Val Val Ala Ile Lys Lys Met Ser Tyr Ser Gly Lys Gln Thr His Glu
50 55 60

Lys Trp Gln Asp Ile Leu Lys Glu Val Lys Phe Leu Arg Gln Leu Lys
65 70 75 80

His Pro Asn Thr Ile Glu Tyr Lys Gly Cys Tyr Leu Lys Glu His Thr
85 90 95

Ala Trp Leu Val Met Glu Tyr Cys Leu Gly Ser Ala Ser Asp Leu Leu
100 105 110

Glu Val His Lys Lys Pro Leu Gln Glu Val Glu Ile Ala Ala Ile Thr
115 120 125

His Gly Ala Leu His Gly Leu Ala Tyr Leu His Ser His Ala Leu Ile
130 135 140

His Arg Asp Ile Lys Ala Gly Asn Ile Leu Leu Thr Glu Pro Gly Gln
145 150 155 160

Val Lys Leu Ala Asp Phe Gly Ser Ala Ser Met Ala Ser Pro Ala Asn
165 170 175

Ser Phe Val Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Ile Leu Ala
180 185 190

Met Asp Glu Gly Gln Tyr Asp Gly Lys Val Asp Ile Trp Ser Leu Gly
195 200 205

Ile Thr Cys Ile Glu Leu Ala Glu Arg Lys Pro Pro Leu Phe Asn Met
210 215 220

Asn Ala Met Ser Ala Leu Tyr His Ile Ala Gln Asn Asp Ser Pro Thr
225 230 235 240

Leu Gln Ser Asn Glu Trp Thr Asp Ser Phe Arg Arg Phe Val Asp Tyr
245 250 255

Cys Leu Gln Lys Ile Pro Gln Glu Arg Pro Thr Ser Ala Glu Leu Leu
260 265 270

Arg His Asp Phe Val Arg Arg Asp Arg Pro Leu Arg Val Leu Ile Asp
275 280 285

Leu Ile Gln Arg Thr Lys Asp Ala Val Arg Glu Leu Asp Asn Leu Gln
290 295 300

Tyr Arg Lys Met Lys Lys Ile Leu Phe Gln Glu Thr Arg Asn Gly Pro
305 310 315 320

Leu Asn Glu Ser Gln Glu Asp Glu Asp Ser Glu His Gly Thr Ser
325 330 335

Leu Asn Arg Glu Met Asp Ser Leu Gly Ser Asn His Ser Ile Pro Ser
340 345 350

Met Ser Val Ser Thr Gly Ser Gln Ser Ser Ser Val Asn Ser Met Gln
355 360 365

Glu Val Met Asp Glu Ser Ser Ser Glu Leu Val Met Met His Asp Asp
370 375 380

Glu Ser Thr Ile Asn Ser Ser Ser Val Val His Lys Lys Asp His
385 390 395 400

Val Phe Thr Arg Asp Glu Ala Gly His Gly Asp Pro Arg Pro Glu Pro
405 410 415

Arg Pro Thr Gln Ser Val Gln Ser Gln Ala Leu His Tyr Arg Asn Arg
420 425 430

Glu Arg Phe Ala Thr Ile Lys Ser Ala Ser Leu Val Thr Arg Gln Ile
435 440 445

His Glu His Glu Gln Glu Asn Glu Leu Arg Glu Gln Met Ser Gly Tyr
450 455 460

Lys Arg Met Arg Arg Gln His Gln Lys Gln Leu Ile Ala Leu Glu Asn
465 470 475 480

Lys Leu Lys Ala Glu Met Asp Glu His Arg Leu Lys Leu Gln Lys Glu
485 490 495

Val Glu Thr His Ala Asn Asn Ser Ser Ile Glu Leu Glu Lys Leu Ala
500 505 510

Lys Lys Gln Val Ala Ile Ile Glu Lys Glu Ala Lys Val Ala Ala Ala
515 520 525

Asp Glu Lys Lys Phe Gln Gln Gln Ile Leu Ala Gln Gln Lys Lys Asp
530 535 540

Leu Thr Thr Phe Leu Glu Ser Gln Lys Lys Gln Tyr Lys Ile Cys Lys
545 550 555 560

Glu Lys Ile Lys Glu Glu Met Asn Glu Asp His Ser Thr Pro Lys Lys

42

565	570	575
Glu Lys Gln Glu Arg Ile Ser Lys His Lys Glu Asn Leu Gln His Thr		
580	585	590
Gln Ala Glu Glu Glu Ala His Leu Leu Thr Gln Gln Arg Leu Tyr Tyr		
595	600	605
Asp Lys Asn Cys Arg Phe Phe Lys Arg Lys Ile Met Ile Lys Arg His		
610	615	620
Glu Val Glu Gln Gln Asn Ile Arg Glu Glu Leu Asn Lys Lys Arg Thr		
625	630	635
640		
Gln Lys Glu Met Glu His Ala Met Leu Ile Arg His Asp Glu Ser Thr		
645	650	655
Arg Glu Leu Glu Tyr Arg Gln Leu His Thr Leu Gln Lys Leu Arg Met		
660	665	670
Asp Leu Ile Arg Leu Gln His Gln Thr Glu Leu Glu Asn Gln Leu Glu		
675	680	685
Tyr Asn Lys Arg Arg Glu Arg Glu Leu His Arg Lys His Val Met Gly		
690	695	700
Leu Arg Gln Gln Pro Lys Asn Leu Lys Ala Met Glu Met Gln Ile Lys		
705	710	715
720		
Lys Gln Phe Gln Asp Thr Cys Lys Val Gln Thr Lys Gln Tyr Lys Ala		
725	730	735
Leu Lys Asn His Gln Leu Glu Val Thr Pro Lys Asn Glu His Lys Thr		
740	745	750
Ile Leu Lys Thr Leu Lys Asp Glu Gln Thr Arg Lys Leu Ala Ile Leu		
755	760	765
Ala Glu Gln Tyr Glu Gln Ser Ile Asn Glu Met Met Ala Ser Gln Ala		
770	775	780
Leu Arg Leu Asp Glu Ala Gln Glu Ala Glu Cys Gln Ala Leu Arg Leu		
785	790	795
800		
Gln Leu Gln Gln Glu Met Glu Leu Leu Asn Ala Tyr Gln Ser Lys Ile		
805	810	815
Lys Met Gln Thr Glu Ala Gln His Glu Arg Glu Leu Gln Lys Leu Glu		
820	825	830
Gln Arg Val Ser Leu Arg Arg Ala His Leu Glu Gln Lys Ile Glu Glu		
835	840	845
Glu Leu Ala Ala Leu Gln Lys Glu Arg Ser Glu Arg Ile Lys Asn Leu		
850	855	860

Leu Glu Arg Gln Glu Arg Glu Ile Glu Thr Phe Asp Met Glu Ser Leu
865 870 875 880

Arg Met Gly Phe Gly Asn Leu Val Thr Leu Asp Phe Pro Lys Glu Asp
885 890 895

Tyr Arg

<210> 23

<211> 786

<212> PRT

<213> Mammalian (Human) SULU3

<400> 23

Ile Glu Leu Ala Glu Arg Lys Pro Pro Leu Phe Asn Met Asn Ala Met
1 5 10 15

Ser Ala Leu Tyr His Ile Ala Gln Asn Glu Ser Pro Thr Leu Gln Ser
20 25 30

Asn Glu Trp Ser Asp Tyr Phe Arg Asn Phe Val Asp Ser Cys Leu Gln
35 40 45

Lys Ile Pro Gln Asp Arg Pro Thr Ser Glu Glu Leu Leu Lys His Ile
50 55 60

Phe Val Leu Arg Glu Arg Pro Glu Thr Val Leu Ile Asp Leu Ile Gln
65 70 75 80

Arg Thr Lys Asp Ala Val Arg Glu Leu Asp Asn Leu Gln Tyr Arg Lys
85 90 95

Met Lys Lys Leu Leu Phe Gln Glu Ala His Asn Gly Pro Ala Val Glu
100 105 110

Ala Gln Glu Glu Glu Glu Gln Asp His Gly Val Gly Arg Thr Gly
115 120 125

Thr Val Asn Ser Val Gly Ser Asn Gln Ser Ile Pro Ser Met Ser Ile
130 135 140

Ser Ala Ser Ser Gln Ser Ser Ser Val Asn Ser Leu Pro Asp Val Ser
145 150 155 160

Asp Asp Lys Ser Glu Leu Asp Met Met Glu Gly Asp His Thr Val Met
165 170 175

Ser Asn Ser Ser Val Ile His Leu Lys Pro Gln Glu Glu Asn Tyr Arg
180 185 190

44

Glu Glu Gly Asp Pro Arg Thr Arg Ala Ser Asp Pro Gln Ser Pro Pro
195 200 205

Gln Val Ser Arg His Lys Ser His Tyr Arg Asn Arg Glu His Phe Ala
210 215 220

Thr Ile Arg Thr Ala Ser Leu Val Thr Arg Gln Met Gln Glu His Glu
225 230 235 240

Gln Asp Ser Glu Leu Arg Glu Gln Met Ser Gly Tyr Lys Arg Met Arg
245 250 255

Arg Gln His Gln Lys Gln Leu Met Thr Leu Glu Asn Lys Leu Lys Ala
260 265 270

Glu Met Asp Glu His Arg Leu Arg Leu Asp Lys Asp Leu Glu Thr Gln
275 280 285

Arg Asn Asn Phe Ala Ala Glu Met Glu Lys Leu Ile Lys Lys His Gln
290 295 300

Ala Ala Met Glu Lys Glu Ala Lys Val Met Ser Asn Glu Glu Lys Lys
305 310 315 320

Phe Gln Gln His Ile Gln Ala Gln Lys Lys Glu Leu Asn Ser Phe
325 330 335

Leu Glu Ser Gln Lys Arg Glu Tyr Lys Leu Arg Lys Glu Gln Leu Lys
340 345 350

Glu Glu Leu Asn Glu Asn Gln Ser Thr Pro Lys Lys Glu Lys Gln Glu
355 360 365

Trp Leu Ser Lys Gln Lys Glu Asn Ile Gln His Phe Gln Ala Glu Glu
370 375 380

Glu Ala Asn Leu Leu Arg Arg Gln Arg Gln Tyr Leu Glu Leu Glu Cys
385 390 395 400

Arg Arg Phe Lys Arg Arg Met Leu Leu Gly Arg His Asn Leu Glu Gln
405 410 415

Asp Leu Val Arg Glu Glu Leu Asn Lys Arg Gln Thr Gln Lys Asp Leu
420 425 430

Glu His Ala Met Leu Leu Arg Gln His Glu Ser Met Gln Glu Leu Glu
435 440 445

Phe Arg His Leu Asn Thr Ile Gln Lys Met Arg Cys Glu Leu Ile Arg
450 455 460

Leu Gln His Gln Thr Glu Leu Thr Asn Gln Leu Glu Tyr Asn Lys Arg
465 470 475 480

Arg Glu Arg Glu Leu Arg Arg Lys His Val Met Glu Val Arg Gln Gln

45

485

490

495

Pro Lys Ser Leu Lys Ser Lys Glu Leu Gln Ile Lys Lys Gln Phe Gln
500 505 510

Asp Thr Cys Lys Ile Gln Thr Arg Gln Tyr Lys Ala Leu Arg Asn His
515 520 525

Leu Leu Glu Thr Thr Pro Lys Ser Glu His Lys Ala Val Leu Lys Arg
530 535 540

Leu Lys Glu Glu Gln Thr Arg Lys Leu Ala Ile Leu Ala Glu Gln Tyr
545 550 555 560

Asp His Ser Ile Asn Glu Met Leu Ser Thr Gln Ala Leu Arg Leu Asp
565 570 575

Glu Ala Gln Glu Ala Glu Cys Gln Val Leu Lys Met Gln Leu Gln Gln
580 585 590

Glu Leu Glu Leu Leu Asn Ala Tyr Gln Ser Lys Ile Lys Met Gln Ala
595 600 605

Glu Ala Gln His Asp Arg Glu Leu Arg Glu Leu Glu Gln Arg Val Ser
610 615 620

Leu Arg Arg Ala Leu Leu Glu Gln Lys Ile Glu Glu Glu Met Leu Ala
625 630 635 640

Leu Gln Asn Glu Arg Thr Glu Arg Ile Arg Ser Leu Leu Glu Arg Gln
645 650 655

Ala Arg Glu Ile Glu Ala Phe Asp Ser Glu Ser Met Arg Leu Gly Phe
660 665 670

Ser Asn Met Val Leu Ser Asn Leu Ser Pro Glu Ala Phe Ser His Ser
675 680 685

Tyr Pro Gly Ala Ser Gly Trp Ser His Asn Pro Thr Gly Gly Pro Gly
690 695 700

Pro His Trp Gly His Pro Met Gly Gly Pro Pro Gln Ala Trp Gly His
705 710 715 720

Pro Met Gln Gly Gly Pro Gln Pro Trp Gly His Pro Ser Gly Pro Met
725 730 735

Gln Gly Val Pro Arg Gly Ser Ser Met Gly Val Arg Asn Ser Pro Gln
740 745 750

Ala Leu Arg Arg Thr Ala Ser Gly Gly Arg Thr Glu Gln Gly Met Ser
755 760 765

Arg Ser Thr Ser Val Thr Ser Gln Ile Ser Asn Gly Ser His Met Ser
770 775 780

Tyr Thr
785

<210> 24
<211> 748
<212> PRT
<213> Mammalian (Murine) SULU3

<400> 24

Met Pro Ser Thr Asn Arg Ala Gly Ser Leu Lys Asp Pro Glu Ile Ala
1 5 10 15

Glu Leu Phe Phe Lys Glu Asp Pro Glu Lys Leu Phe Thr Asp Leu Arg
20 25 30

Glu Ile Gly His Gly Ser Phe Gly Ala Val Tyr Phe Ala Arg Asp Val
35 40 45

Arg Thr Asn Glu Val Val Ala Ile Lys Lys Met Ser Tyr Ser Gly Lys
50 55 60

Gln Ser Thr Glu Lys Trp Gln Asp Ile Ile Lys Glu Val Lys Phe Leu
65 70 75 80

Gln Arg Ile Lys His Pro Asn Ser Ile Glu Tyr Lys Gly Cys Tyr Leu
85 90 95

Arg Glu His Thr Ala Trp Leu Val Met Glu Tyr Cys Leu Gly Ser Ala
100 105 110

Ser Asp Leu Leu Glu Val His Lys Lys Pro Leu Gln Glu Val Glu Ile
115 120 125

Ala Ala Ile Thr His Gly Ala Leu Gln Gly Leu Ala Tyr Leu His Ser
130 135 140

His Thr Met Ile His Arg Asp Ile Lys Ala Gly Asn Ile Leu Leu Thr
145 150 155 160

Glu Pro Gly Gln Val Lys Leu Ala Asp Phe Gly Ser Ala Ser Met Ala
165 170 175

Ser Pro Ala Asn Ser Phe Val Gly Thr Pro Tyr Trp Met Ala Pro Glu
180 185 190

Val Ile Leu Ala Met Asp Glu Gly Gln Tyr Asp Gly Lys Val Asp Val
195 200 205

Trp Ser Leu Gly Ile Thr Cys Ile Glu Leu Ala Glu Arg Lys Pro Pro
210 215 220

Leu Phe Asn Met Asn Ala Met Ser Ala Leu Tyr His Ile Ala Gln Asn
 225 230 235 240

 Glu Ser Pro Thr Leu Gln Ser Asn Met Asn Asp Ser Cys Leu Gln Lys
 245 250 255

 Ile Pro Gln Asp Arg Pro Thr Ser Glu Glu Leu Leu Lys His Met Phe
 260 265 270

 Val Leu Arg Glu Arg Pro Glu Thr Val Leu Ile Asp Leu Ile Gln Arg
 275 280 285

 Thr Lys Asp Ala Val Arg Glu Leu Asp Asn Leu Gln Tyr Arg Lys Met
 290 295 300

 Lys Lys Leu Leu Phe Gln Glu Ala His Asn Gly Pro Ala Val Glu Ala
 305 310 315 320

 Gln Glu Glu Glu Glu Gln Asp His Gly Val Gly Arg Thr Gly Thr
 325 330 335

 Val Asn Ser Val Gly Ser Asn Gln Ser Ile Pro Ser Met Ser Ile Ser
 340 345 350

 Ala Ser Ser Gln Ser Ser Val Asn Ser Leu Pro Asp Ala Ser Asp
 355 360 365

 Asp Lys Ser Glu Leu Asp Met Met Glu Gly Asp His Thr Val Met Ser
 370 375 380

 Asn Ser Ser Val Ile His Leu Lys Pro Glu Glu Asn Tyr Gln Glu
 385 390 395 400

 Glu Gly Asp Pro Arg Thr Arg Ala Ser Asp Pro Gln Ser Pro Pro Gln
 405 410 415

 Val Ser Arg His Lys Ser His Tyr Arg Asn Arg Glu His Phe Ala Thr
 420 425 430

 Ile Arg Thr Ala Ser Leu Val Thr Arg Gln Met Gln Glu His Glu Gln
 435 440 445

 Asp Ser Glu Leu Arg Glu Gln Met Ser Gly Tyr Lys Arg Met Arg Arg
 450 455 460

 Gln His Gln Lys Gln Leu Met Thr Leu Glu Asn Lys Leu Lys Ala Glu
 465 470 475 480

 Met Asp Glu His Arg Leu Arg Leu Asp Lys Asp Leu Glu Thr Gln Arg
 485 490 495

 Asn Asn Phe Ala Ala Glu Met Glu Lys Leu Ile Lys Lys His Gln Ala
 500 505 510

 Ala Met Glu Lys Glu Ala Lys Val Met Ala Asn Glu Glu Lys Lys Phe

515	520	525
-----	-----	-----

Gln Gln His Ile Gln Ala Gln Gln Lys Lys Glu Leu Asn Ser Phe Leu		
530	535	540

Glu Ser Gln Lys Arg Glu Tyr Lys Leu Arg Lys Glu Gln Leu Lys Glu		
545	550	560

Glu Leu Asn Glu Asn Gln Ser Thr Pro Lys Lys Glu Lys Gln Glu Trp		
565	570	575

Leu Ser Lys Gln Lys Glu Asn Ile Gln His Phe Gln Ala Glu Glu Glu		
580	585	590

Ala Asn Leu Leu Arg Arg Gln Arg Gln Tyr Leu Glu Leu Glu Cys Arg		
595	600	605

Arg Phe Lys Arg Arg Met Leu Leu Gly Arg His Asn Leu Glu Gln Asp		
610	615	620

Leu Val Arg Glu Glu Leu Asn Lys Arg Gln Thr Gln Lys Asp Leu Glu		
625	630	640

His Ala Met Leu Leu Arg Gln His Glu Ser Met Gln Glu Leu Glu Phe		
645	650	655

Arg His Leu Asn Thr Ile Gln Lys Met Arg Cys Glu Leu Ile Arg Leu		
660	665	670

Gln His Gln Thr Glu Leu Thr Asn Gln Leu Glu Tyr Asn Lys Arg Arg		
675	680	685

Glu Arg Glu Leu Arg Arg Lys His Val Met Glu Val Arg Gln Gln Pro		
690	695	700

Lys Ser Leu Lys Ser Lys Glu Leu Gln Ile Lys Lys Gln Phe Gln Asp		
705	710	720

Thr Cys Lys Ile Gln Thr Arg Gln Tyr Lys Ala Leu Arg Asn His Leu		
725	730	735

Leu Glu Thr Thr Pro Lys Asn Glu His Lys Ala Ile		
740	745	

<210> 25

<211> 2795

<212> DNA

<213> Mammalian (Human) GEK2

<400> 25

cgaagccaca gcccggagcc gagccccgagc ccgagccggc gccaccgcgc ccccgccat	60
ggctttgcc aatttccgcc gcatactgcg cctgtctacc ttcgagaaga gaaagtcccg	120

cgaatatacgac cacgtccggcc	gcccacactggaa	cccccaacggag	gtgtgggaga	tcgtgggcga	180
gctggggcggac	ggcgcccttcg	gcaagggttta	caaggccaaag	aataaggaga	240
ggctgcggcc	aaagtcatcg	aaaccaggaaag	tgaggaggag	ctggaggact	300
gattgagatc	ctggccaccc	gcccaccc	ctacattgt	aagctcctgg	360
tcacgacggg	aagctgttgc	tcatgattga	gttctgtcca	gggggagccg	420
catgctggag	ctggacagaaag	gcctcacgg	gccccagata	caggtggttt	480
gctagaagcc	ctcaacttcc	tgcacagcaa	gaggatcatc	caccgagatc	540
caacgtgctg	atgaccctcg	agggagacat	caggtggct	gactttgttgc	600
aatctgaag	actctacaga	aacgagattc	cttcatcg	acgccttact	660
cgaggtggtc	atgtgtgaga	ccatgaaaaga	cacgccc	gactacaaag	720
gtccctggc	atcacgctg	ttgagatggc	ccagatcg	ccgcccacacc	780
ccccatgcgg	gtcctgtctaa	agatccaa	gtcggaccc	cccacgctgc	840
caagtggct	gttagagttcc	gtgacttc	gaagatagcc	ctggataaga	900
ccgaccagg	gccgcgcage	tgctggagca	tcccttcgtc	agcagcatca	960
ggctctgcgg	gagctgggttgc	ctgaggccaa	ggccgaggtg	atgaaagaga	1020
ccgggatgag	ggggaaagagg	aggacgcgt	ggatgccc	tccaccctgg	1080
tcagaactcc	tctgaggtg	gtccgccaag	cctcaatgt	gacaagc	1140
accttccacc	ccgctggcac	ccagccagtc	tcaggacagt	gtgaatgagc	1200
gccctctggg	gacagatccc	tccaaaccac	cagtcccc	ctcg	1260
gaacggcctg	gcagtgcc	tgccc	gaagtcc	ggcc	1320
aattcaggta	gcccaggaga	agcaagttgc	tgagcagg	ggggac	1380
caacagatct	caaaaggcc	gccagagcc	gccc	acagc	1440
tggggagaag	ctggccaatg	gcagcctg	gccac	ctgg	1500
caagaggac	tccgactg	gcagc	cac	ggc	1560
tctctccact	gac	ctgtc	gat	gg	1620
actgtacaaa	aaaacc	caaa	ttgt	gg	1680
catcaccacc	tccaa	accac	gtt	gg	1740
caggcgc	ccag	ttgt	gg	gg	1800
ccagctg	actcc	cc	gg	gg	1860
aatcaacg	ccat	ccat	gg	gg	1920
gcagca	aaat	aaat	gg	gg	1980
gagaagatgg	ttgt	ttgt	gg	gg	2040
gatccgc	ccat	ccat	gg	gg	2100
gaaagaggt	ttgt	ttgt	gg	gg	2160
gcagaagat	gg	gg	gg	gg	2220
gac	gg	gg	gg	gg	2280
ctgt	gg	gg	gg	gg	2340
gac	gg	gg	gg	gg	2400
gg	gg	gg	gg	gg	2460
gg	gg	gg	gg	gg	2520
caagaag	gg	gg	gg	gg	2580
gcag	gg	gg	gg	gg	2640
ttct	gg	gg	gg	gg	2700
cc	gg	gg	gg	gg	2760
ctggat	gg	gg	gg	gg	2795

<210> 26

<211> 912

<212> PRT

<213> Mammalian (Human) GEK2

<400> 26

Met Ala Phe Ala Asn Phe Arg Arg Ile Leu Arg Leu Ser Thr Phe Glu

50
1 5 10 15

Lys Arg Lys Ser Arg Glu Tyr Glu His Val Arg Arg Asp Leu Asp Pro
20 25 30

Asn Glu Val Trp Glu Ile Val Gly Glu Leu Gly Asp Gly Ala Phe Gly
35 40 45

Lys Val Tyr Lys Ala Lys Asn Lys Glu Thr Gly Ala Leu Ala Ala Ala
50 55 60

Lys Val Ile Glu Thr Lys Ser Glu Glu Leu Glu Asp Tyr Ile Val
65 70 75 80

Glu Ile Glu Ile Leu Ala Thr Cys Asp His Pro Tyr Ile Val Lys Leu
85 90 95

Leu Gly Ala Tyr Tyr His Asp Gly Lys Leu Trp Ile Met Ile Glu Phe
100 105 110

Cys Pro Gly Gly Ala Val Asp Ala Ile Met Leu Glu Leu Asp Arg Gly
115 120 125

Leu Thr Glu Pro Gln Ile Gln Val Val Cys Arg Gln Met Leu Glu Ala
130 135 140

Leu Asn Phe Leu His Ser Lys Arg Ile Ile His Arg Asp Leu Lys Ala
145 150 155 160

Gly Asn Val Leu Met Thr Leu Glu Gly Asp Ile Arg Leu Ala Asp Phe
165 170 175

Gly Val Ser Ala Lys Asn Leu Lys Thr Leu Gln Lys Arg Asp Ser Phe
180 185 190

Ile Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Val Met Cys Glu Thr
195 200 205

Met Lys Asp Thr Pro Tyr Asp Tyr Lys Ala Asp Ile Trp Ser Leu Gly
210 215 220

Ile Thr Leu Ile Glu Met Ala Gln Ile Glu Pro Pro His His Glu Leu
225 230 235 240

Asn Pro Met Arg Val Leu Leu Lys Ile Ala Lys Ser Asp Pro Pro Thr
245 250 255

Leu Leu Thr Pro Ser Lys Trp Ser Val Glu Phe Arg Asp Phe Leu Lys
260 265 270

Ile Ala Leu Asp Lys Asn Pro Glu Thr Arg Pro Ser Ala Ala Gln Leu
275 280 285

Leu Glu His Pro Phe Val Ser Ser Ile Thr Ser Asn Lys Ala Leu Arg
290 295 300

Glu Leu Val Ala Glu Ala Lys Ala Glu Val Met Glu Glu Ile Glu Asp
 305 310 315 320

Gly Arg Asp Glu Gly Glu Glu Asp Ala Val Asp Ala Ala Ser Thr
 325 330 335

Leu Glu Asn His Thr Gln Asn Ser Ser Glu Val Ser Pro Pro Ser Leu
 340 345 350

Asn Ala Asp Lys Pro Leu Glu Glu Ser Pro Ser Thr Pro Leu Ala Pro
 355 360 365

Ser Gln Ser Gln Asp Ser Val Asn Glu Pro Cys Ser Gln Pro Ser Gly
 370 375 380

Asp Arg Ser Leu Gln Thr Thr Ser Pro Pro Val Val Ala Pro Gly Asn
 385 390 395 400

Glu Asn Gly Leu Ala Val Pro Val Pro Leu Arg Lys Ser Arg Pro Val
 405 410 415

Ser Met Asp Ala Arg Ile Gln Val Ala Gln Glu Lys Gln Val Ala Glu
 420 425 430

Gln Gly Gly Asp Leu Ser Pro Ala Ala Asn Arg Ser Gln Lys Ala Ser
 435 440 445

Gln Ser Arg Pro Asn Ser Ser Ala Leu Glu Thr Leu Gly Gly Glu Lys
 450 455 460

Leu Ala Asn Gly Ser Leu Glu Pro Pro Ala Gln Ala Ala Pro Gly Pro
 465 470 475 480

Ser Lys Arg Asp Ser Asp Cys Ser Ser Leu Cys Thr Ser Glu Ser Met
 485 490 495

Asp Tyr Gly Thr Asn Leu Ser Thr Asp Leu Ser Leu Asn Lys Glu Met
 500 505 510

Gly Ser Leu Ser Ile Lys Asp Pro Lys Leu Tyr Lys Lys Thr Leu Lys
 515 520 525

Arg Thr Arg Lys Phe Val Val Asp Gly Val Glu Val Ser Ile Thr Thr
 530 535 540

Ser Lys Ile Ile Ser Glu Asp Glu Lys Lys Asp Glu Glu Met Arg Phe
 545 550 555 560

Leu Arg Arg Gln Glu Leu Arg Glu Leu Arg Leu Leu Gln Lys Glu Glu
 565 570 575

His Arg Asn Gln Thr Gln Leu Ser Asn Lys His Glu Leu Gln Leu Glu
 580 585 590

Gln Met His Lys Arg Phe Glu Gln Glu Ile Asn Ala Lys Lys Lys Phe
 595 600 605

Phe Asp Thr Glu Leu Glu Asn Leu Glu Arg Gln Gln Lys Gln Gln Val
610 615 620

Glu Lys Met Glu Gln Asp His Ala Val Arg Arg Arg Glu Glu Ala Arg
625 630 635 640

Arg Ile Arg Leu Glu Gln Asp Arg Asp Tyr Thr Arg Phe Gln Glu Gln
645 650 655

Leu Lys Leu Met Lys Lys Glu Val Lys Asn Glu Val Glu Lys Leu Pro
660 665 670

Arg Gln Gln Arg Lys Glu Ser Met Lys Gln Lys Met Glu Glu His Thr
675 680 685

Gln Lys Lys Gln Leu Leu Asp Arg Asp Phe Val Ala Lys Gln Lys Glu
690 695 700

Asp Leu Glu Leu Ala Met Lys Arg Leu Thr Thr Asp Asn Arg Arg Glu
705 710 715 720

Ile Cys Asp Lys Glu Arg Glu Cys Leu Met Lys Lys Gln Glu Leu Leu
725 730 735

Arg Asp Arg Glu Ala Ala Leu Trp Glu Met Glu Glu His Gln Leu Gln
740 745 750

Glu Arg His Gln Leu Val Lys Gln Gln Leu Lys Asp Gln Tyr Phe Leu
755 760 765

Gln Arg His Glu Leu Leu Arg Lys His Glu Lys Glu Arg Glu Gln Met
770 775 780

Gln Arg Tyr Asn Gln Arg Met Ile Glu Gln Leu Lys Val Arg Gln Gln
785 790 795 800

Gln Glu Lys Ala Arg Leu Pro Lys Ile Gln Arg Ser Glu Gly Lys Thr
805 810 815

Arg Met Ala Met Tyr Lys Ser Leu His Ile Asn Gly Gly Ser
820 825 830

Ala Ala Glu Gln Arg Glu Lys Ile Lys Gln Phe Ser Gln Gln Glu Glu
835 840 845

Lys Arg Gln Lys Ser Glu Arg Leu Gln Gln Gln Lys His Glu Asn
850 855 860

Gln Met Arg Asp Met Leu Ala Gln Cys Glu Ser Asn Met Ser Glu Leu
865 870 875 880

Gln Gln Leu Gln Asn Glu Lys Cys His Leu Leu Val Glu His Glu Thr
885 890 895

Gln Lys Leu Lys Ala Leu Asp Glu Ser His Asn Gln Asn Leu Lys Glu
900 905 910

<210> 27
<211> 3604
<212> DNA
<213> Mammalian (Human) PAK4

<400> 27

cggtttccggg	cttcccgcctc	cgcaggcccg	cgaggactg	gcccaagcaag	gtccccaggc	60
ttccctctcc	tttaggcgccta	agagagaggc	ccagtgcggg	tgaggagtcg	cgaggaagag	120
gcggaaaggcg	ccggaaaggca	ccatgttccg	caaaaaaaag	aagaaacgc	ctgagatctc	180
agcgccacag	aacctccagc	accgtgtcca	caccccttc	gacccaaag	aaggcaagtt	240
tgtgggcctc	cccccacaat	ggcagaacat	cctggacaca	ctgcggcgcc	ccaagcccg	300
ggtggacccct	tgcgaatca	cacgggtgca	gctccagccc	atgaagacag	tggtgccggg	360
cagcgcgatg	cctgtggatg	gctacatetc	ggggctgctc	aacgacatcc	agaagttgtc	420
agtcatcagc	tccaacaccc	tgcgtggccg	cagccccacc	agccggcgcc	gggcacagtc	480
cctggggctg	ctgggggatg	agcactgggc	caccgaccca	gacatgtacc	tccagagccc	540
ccagtcgtag	cgcactgacc	cccacggcct	ctacctcagc	tgcaacgggg	gcacaccagc	600
aggccacaag	cagatgccgt	ggcccgagcc	acagagccca	cgggtcctgc	ccaatggct	660
ggctgcaaag	gcacagtccc	tggggcccgc	cgagtttcag	ggtgcctcgc	agcgctgtct	720
gcagctgggt	gcctgcctgc	agagctcccc	accaggagcc	tcgcccccca	cgggcaccaa	780
taggcatgga	atgaaggctg	ccaagcatgg	ctctgaggag	gcccgccac	agtctgcct	840
ggtgggctca	gccacaggca	ggccaggtgg	ggaaggcagc	cctagcccta	agaccggga	900
gagcagccctg	aagcgcagggc	tattccgaaag	catgttccctg	tccactgtctg	ccacagcccc	960
tccaagcagc	agcaagccag	gccctccacc	acagagcaag	cccaactcct	ctttccgacc	1020
gcccgagaaa	gacaacccccc	caaggcttgt	ggccaaggcc	cagtccctgc	cctcggacca	1080
qccgggtgggg	accttcagcc	ctctgaccac	ttcggatacc	agcagcccccc	agaagtcct	1140
ccgcacagcc	ccggccacag	gcagcttcc	aggccggct	tccccagcgg	gatccccccg	1200
cacctgcac	gcccgatca	gcaccagcaa	cctgtacctg	ccccaggacc	ccacgggtgc	1260
caaggggtgcc	ctggctgggt	aggacacagg	tgttgtgaca	catgagcagt	tcaaggctgc	1320
gctcagatg	gtggtgggacc	agggtgacc	ccggctgtct	ctggacagct	acgtgaagat	1380
tggcgagggc	tccaccggca	tcgtctgctt	ggcccgggag	aagcactcgg	gccgcaggt	1440
ggccgtcaag	atgatggacc	tcaggaagca	gcagcgcagg	gagctgctct	tcaacgaggt	1500
ggtgatcatg	cgggactacc	agcacttcaa	cgtggtgag	atgtacaaga	gctacctggt	1560
gggcgaggag	ctgtgggtgc	tcatggagtt	cctgcaggga	ggagccctca	cagacatcgt	1620
ctcccaagtc	aggctgaatg	aggagcaat	tgccactgtg	tgtgaggctg	tgctgcaggc	1680
cctggcttac	ctgcgtgtctc	aggggtgtcat	ccaccggac	atcaagagt	actccatcct	1740
gctgaccctc	gatggcaggg	tgaagctctc	ggacttcgga	ttctgtgctc	agatcagcaa	1800
agacgtccct	aagaggaagt	ccctgggg	aacccctac	tggatggctc	ctgaagtgt	1860
ctccaggctt	ttgttatgcca	ctgaggtgga	tatctggct	ctgggcatca	tggtgattga	1920
gatggtagat	ggggagccac	cgtacttcag	tgactccca	gtgcaagcca	tgaagaggct	1980
ccgggacacgc	cccccaccca	agctaaaaaa	ctctcacaag	gtctcccaag	tgctgcgaga	2040
cttcctggag	cggatgtcg	tgcgggaccc	ccaagagaga	gccacagccc	aggacgtctct	2100
agaccacccc	ttctctgtgc	agacagggt	acctgagtgc	ctggtgcccc	tgatccagct	2160
ctaccgaaag	cagacctcca	cctgtgtgac	ccaccccaag	tatgcctgcc	acctacgccc	2220
acaggcaggg	cacactgggc	agccagccctg	ccggcaggac	ttgcctgcct	cctccctctca	2280
gtattctctc	caaagattga	aatgtgaagc	cccagcccca	ccctctgccc	ttcagctac	2340
tgggccaggc	cggacctgccc	ccctcagtgt	ctctccctcc	cgagtcccca	gatggagacc	2400
cctttctaca	ggatgacccc	ttgatatttg	cacaggata	tttctaagaa	acgcagaggc	2460
cagcgttccct	ggcctctgca	gccaacacag	tagaaaaggc	tgcgtgtggtt	ttttaaaggc	2520
agttgtccac	tagtgcctca	ggccactgca	gagggcagac	tgcgtgtctc	cacaqatacc	2580

tgctgttctc agctccagct tcaaaccctcg	agtctcgaga gggccacggg	gtggtttta	2640
tgaccggaaat cccgcttcct ccctcacgtc	tgtatgtcctg aagggtgcagt	cccacctgt	2700
cagccccctcc ccgccaagaa ctgtqaatgg	cctgctccag gccatggctg	ggggcaggga	2760
gtgagggggac aatttcttag	agaatgggtt	cggtgtgaa	2820
ctttacagaa tggagagaac atcgtgtgt	tgtgtgtgt	tgtgtgtgt	2880
tgtgtgtgt	tgtgtgtgt	tgtgtgtgt	2940
cctccaggtc acccacacgcc agttcagga	gggagggaaag	ccacccctgac	2940
cctgaaggga cctgctttct tggcctggct	aggctcccc	tctctccac	3000
cagtggagtg gtccatgtca	acttaaacac	tccacctctc	3060
tcctgtactg tgcatactgtga	gggcccagaga	cactcctgtg	3120
cattcaagtt tatctcaata accaggcat	caggaatgt	tctacactgc	3180
tctgggagag aaggcgaacc acctccgggt	aggattggca	tctgagcttc	3180
gtgtgaaac catctcccc	actgtgcctcc	actgtgttac	3240
cctgccccac ttcaagggtc	ctgtgagatg	caggaccca	3240
accctttcta tctagataat atttttagat	gaattccagg	ttgttctt	3300
caaagaaatt gcaaggactt ttttaagg	aaagaacttc	cagaacccta	3360
ctagaaattt ttgtgaattt tttgcacttg	agggtctgg	3420	
agcc	tgcctgttt	taaaaacaaa	3480
	aaattaaatt	agcatcttcc	3540
	gagtgttcaa	3600	
		3604	

<210> 28
<211> 2050
<212> DNA
<213> Mammalian (Human) PAK5

<400> 28

ggccagtggg gcgaaactgg	cagctggccg	gcccttaac	acctacccga	gggctgacac	60
ggaccaccca tccccgggtg	cccaggggga	gcctcatgac	gtggccctta	acggggccatc	120
agcggggggc ctggccatcc	cccagtccctc	ctccctcctcc	tcccggcctc	ccacccgagc	180
ccgaggtgcc	cccagccctg	gagtgtggg	accccaacgcc	tcagagcccc	240
tccagccctgc	accccccggc	cccctgtgt	tcctgggccc	cctggccccc	300
gcgggagcca	cagcgagtat	cccatgagca	gttccggct	gcctgcagc	360
cccaggcgac	ccccgctct	acctggacaa	cttcatcaag	ttggcgagg	420
catcggtgc	atcgccaccc	tgcgcagtc	gggcaagctg	gtggccgtca	480
cctgcgcgaag	cagcagaggc	gcfagctgt	cttcaacgag	gtggtaatca	540
ccagcacgag	aatgtgggg	agatgtacaa	cagctaccc	tggtgggacg	600
gtcatggag	ttcctggaaag	gagggccct	caccgacatc	gtcaccacaca	660
cgaggagcag	atcgccggcc	tgtgccttc	agtgcgtcag	gccctgtcgg	720
ccaggcgctc	atccacccgg	acatcaagag	cgactcgatc	tgctgtacccc	780
ggtgaagctg	tcagacttt	gttctgcgc	ccaggtgagc	aaggaaatgc	840
gtcgctggtc	ggcacgcct	actggatggc	cccagagctc	cccgaaaggaa	900
gccagaggta	gacatctgtt	cgctgggat	aatgggtatt	gagatgggg	960
cccctacttc	aacgagccac	ccctcaaagc	catgaagatg	acggagagcc	1020
ccgactgaag	aacctgcaca	aggtgtgc	atccctgaag	ttccctacgg	1080
ggtgcgagac	cctgcccagc	ggggccacggc	cccgagactc	atctcccgcc	1140
caaggcgagg	ccgcctgca	gcatcggtc	cctcatgcgc	cataaccgc	1200
cccagcgccc	ttccctcaa	ccaaagagcc	ccccgggtc	ccatggggcc	1260
agttagggggc	caggcctccc	actcctccca	ccccgggaga	ggctcccggt	1320
tccttgctgg	gggttagatga	gaccctacta	tgctcccggt	ggcaccaccc	1380
agaaaaaacac	agggactctgt	gggagcaagc	gaggctccca	cctctggac	1440
aggccctccc	ccatgttctt	ctgtctccag	ggagggcagc	ggccctccca	1500
tctgcagtgg	gggtcgctgg	gggtggagag	aacactaaga	tcactggaa	1560
tgcacgcgtg	tgagtgtgca	tgtgtgtgt	tgtgcaaagg	tccagccacc	1620
cccccaagg	ggtgtctggc	gccttgcctg	acacccagcc	ccctctcccc	1680

gtgggggtcg atcatgaatg tccgaagagt ggcctttcc cgtagccctg cgccccctt 1740
 ctgtggctgg atggggagac aggtcagggc ccccccacctt ctccagcccc tgcagcaaat 1800
 gactactgca cctggacacgc ctcctttt ctagaagtct atttatattt tcattttata 1860
 acactcttagc ccctgcctt attggggac agatggtccc tgctctgcgg ggtggccctg 1920
 gcagaaccac tgctgaaga accaggttcc tgcccggtca ggcagcccc agcccgccca 1980
 cccctgcctc gagttagttt tacaattaaa acattgtctt gtttgtgaa aaaaaaaaaa 2040
 aaaaaaaaaaa 2050

<210> 29
 <211> 681
 <212> PRT
 <213> Mammalian (Human) PAK4

<400> 29

Met Phe Arg Lys Lys Lys Lys Arg Pro Glu Ile Ser Ala Pro Gln
 1 5 10 15

Asn Phe Gln His Arg Val His Thr Ser Phe Asp Pro Lys Glu Gly Lys
 20 25 30

Phe Val Gly Leu Pro Pro Gln Trp Gln Asn Ile Leu Asp Thr Leu Arg
 35 40 45

Arg Pro Lys Pro Val Val Asp Pro Ser Arg Ile Thr Arg Val Gln Leu
 50 55 60

Gln Pro Met Lys Thr Val Val Arg Gly Ser Ala Met Pro Val Asp Gly
 65 70 75 80

Tyr Ile Ser Gly Leu Leu Asn Asp Ile Gln Lys Leu Ser Val Ile Ser
 85 90 95

Ser Asn Thr Leu Arg Gly Arg Ser Pro Thr Ser Arg Arg Arg Ala Gln
 100 105 110

Ser Leu Gly Leu Leu Gly Asp Glu His Trp Ala Thr Asp Pro Asp Met
 115 120 125

Tyr Leu Gln Ser Pro Gln Ser Glu Arg Thr Asp Pro His Gly Leu Tyr
 130 135 140

Leu Ser Cys Asn Gly Gly Thr Pro Ala Gly His Lys Gln Met Pro Trp
 145 150 155 160

Pro Glu Pro Gln Ser Pro Arg Val Leu Pro Asn Gly Leu Ala Ala Lys
 165 170 175

Ala Gln Ser Leu Gly Pro Ala Glu Phe Gln Gly Ala Ser Gln Arg Cys
 180 185 190

Leu Gln Leu Gly Ala Cys Leu Gln Ser Ser Pro Pro Gly Ala Ser Pro
 195 200 205

Pro Thr Gly Thr Asn Arg His Gly Met Lys Ala Ala Lys His Gly Ser
210 215 220

Glu Glu Ala Arg Pro Gln Ser Cys Leu Val Gly Ser Ala Thr Gly Arg
225 230 235 240

Pro Gly Gly Glu Gly Ser Pro Ser Pro Lys Thr Arg Glu Ser Ser Leu
245 250 255

Lys Arg Arg Leu Phe Arg Ser Met Phe Leu Ser Thr Ala Ala Thr Ala
260 265 270

Pro Pro Ser Ser Ser Lys Pro Gly Pro Pro Pro Gln Ser Lys Pro Asn
275 280 285

Ser Ser Phe Arg Pro Pro Gln Lys Asp Asn Pro Pro Ser Leu Val Ala
290 295 300

Lys Ala Gln Ser Leu Pro Ser Asp Gln Pro Val Gly Thr Phe Ser Pro
305 310 315 320

Leu Thr Thr Ser Asp Thr Ser Ser Pro Gln Lys Ser Leu Arg Thr Ala
325 330 335

Pro Ala Thr Gly Gln Leu Pro Gly Arg Ser Ser Pro Ala Gly Ser Pro
340 345 350

Arg Thr Trp His Ala Gln Ile Ser Thr Ser Asn Leu Tyr Leu Pro Gln
355 360 365

Asp Pro Thr Val Ala Lys Gly Ala Leu Ala Gly Glu Asp Thr Gly Val
370 375 380

Val Thr His Glu Gln Phe Lys Ala Ala Leu Arg Met Val Val Asp Gln
385 390 395 400

Gly Asp Pro Arg Leu Leu Leu Asp Ser Tyr Val Lys Ile Gly Glu Gly
405 410 415

Ser Thr Gly Ile Val Cys Leu Ala Arg Glu Lys His Ser Gly Arg Gln
420 425 430

Val Ala Val Lys Met Met Asp Leu Arg Lys Gln Gln Arg Arg Glu Leu
435 440 445

Leu Phe Asn Glu Val Val Ile Met Arg Asp Tyr Gln His Phe Asn Val
450 455 460

Val Glu Met Tyr Lys Ser Tyr Leu Val Gly Glu Glu Leu Trp Val Leu
465 470 475 480

Met Glu Phe Leu Gln Gly Gly Ala Leu Thr Asp Ile Val Ser Gln Val
485 490 495

Arg Leu Asn Glu Glu Gln Ile Ala Thr Val Cys Glu Ala Val Leu Gln
500 505 510

Ala Leu Ala Tyr Leu His Ala Gln Gly Val Ile His Arg Asp Ile Lys
 515 520 525

Ser Asp Ser Ile Leu Leu Thr Leu Asp Gly Arg Val Lys Leu Ser Asp
 530 535 540

Phe Gly Phe Cys Ala Gln Ile Ser Lys Asp Val Pro Lys Arg Lys Ser
 545 550 555 560

Leu Val Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Ile Ser Arg Ser
 565 570 575

Leu Tyr Ala Thr Glu Val Asp Ile Trp Ser Leu Gly Ile Met Val Ile
 580 585 590

Glu Met Val Asp Gly Glu Pro Pro Tyr Phe Ser Asp Ser Pro Val Gln
 595 600 605

Ala Met Lys Arg Leu Arg Asp Ser Pro Pro Pro Lys Leu Lys Asn Ser
 610 615 620

His Lys Val Ser Pro Val Leu Arg Asp Phe Leu Glu Arg Met Leu Val
 625 630 635 640

Arg Asp Pro Gln Glu Arg Ala Thr Ala Gln Glu Leu Leu Asp His Pro
 645 650 655

Phe Leu Leu Gln Thr Gly Leu Pro Glu Cys Leu Val Pro Leu Ile Gln
 660 665 670

Leu Tyr Arg Lys Gln Thr Ser Thr Cys
 675 680

<210> 30
 <211> 398
 <212> PRT
 <213> Mammalian (Human) PAK5

<400> 30

Ala Ser Gly Ala Lys Leu Ala Ala Gly Arg Pro Phe Asn Thr Tyr Pro
 1 5 10 15

Arg Ala Asp Thr Asp His Pro Ser Arg Gly Ala Gln Gly Glu Pro His
 20 25 30

Asp Val Ala Pro Asn Gly Pro Ser Ala Gly Gly Leu Ala Ile Pro Gln
 35 40 45

Ser Ser Ser Ser Ser Arg Pro Pro Thr Arg Ala Arg Gly Ala Pro
 50 55 60

58

Ser Pro Gly Val Leu Gly Pro His Ala Ser Glu Pro Gln Leu Ala Pro
65 70 75 80

Pro Ala Cys Thr Pro Ala Ala Pro Ala Val Pro Gly Pro Pro Gly Pro
85 90 95

Arg Ser Pro Gln Arg Glu Pro Gln Arg Val Ser His Glu Gln Phe Arg
100 105 110

Ala Ala Leu Gln Leu Val Val Asp Pro Gly Asp Pro Arg Ser Tyr Leu
115 120 125

Asp Asn Phe Ile Lys Ile Gly Glu Gly Ser Thr Gly Ile Val Cys Ile
130 135 140

Ala Thr Val Arg Ser Ser Gly Lys Leu Val Ala Val Lys Lys Met Asp
145 150 155 160

Leu Arg Lys Gln Gln Arg Arg Glu Leu Leu Phe Asn Glu Val Val Ile
165 170 175

Met Arg Asp Tyr Gln His Glu Asn Val Val Glu Met Tyr Asn Ser Tyr
180 185 190

Leu Val Gly Asp Glu Leu Trp Val Val Met Glu Phe Leu Glu Gly Gly
195 200 205

Ala Leu Thr Asp Ile Val Thr His Thr Arg Met Asn Glu Glu Gln Ile
210 215 220

Ala Ala Val Cys Leu Ala Val Leu Gln Ala Leu Ser Val Leu His Ala
225 230 235 240

Gln Gly Val Ile His Arg Asp Ile Lys Ser Asp Ser Ile Leu Leu Thr
245 250 255

His Asp Gly Arg Val Lys Leu Ser Asp Phe Gly Phe Cys Ala Gln Val
260 265 270

Ser Lys Glu Val Pro Arg Arg Lys Ser Leu Val Gly Thr Pro Tyr Trp
275 280 285

Met Ala Pro Glu Leu Ile Ser Arg Leu Pro Tyr Gly Pro Glu Val Asp
290 295 300

Ile Trp Ser Leu Gly Ile Met Val Ile Glu Met Val Asp Gly Glu Pro
305 310 315 320

Pro Tyr Phe Asn Glu Pro Pro Leu Lys Ala Met Lys Met Ile Arg Asp
325 330 335

Asn Leu Pro Pro Arg Leu Lys Asn Leu His Lys Val Ser Pro Ser Leu
340 345 350

Lys Gly Phe Leu Asp Arg Leu Leu Val Arg Asp Pro Ala Gln Arg Ala
355 360 365

Thr Ala Ala Glu Leu Leu Lys His Pro Phe Leu Ala Lys Ala Gly Pro
370 375 380

Pro Ala Ser Ile Val Pro Leu Met Arg Gln Asn Arg Thr Arg
385 390 395

<210> 31
<211> 1001
<212> PRT
<213> Mammalian (murine/human) SULU3

<400> 31

-Met Pro Ser Thr Asn Arg Ala Gly Ser Leu Lys Asp Pro Glu Ile Ala
1 5 10 15

Glu Leu Phe Phe Lys Glu Asp Pro Glu Lys Leu Phe Thr Asp Leu Arg
20 25 30

Glu Ile Gly His Gly Ser Phe Gly Ala Val Tyr Phe Ala Arg Asp Val
35 40 45

Arg Thr Asn Glu Val Val Ala Ile Lys Lys Met Ser Tyr Ser Gly Lys
50 55 60

Gln Ser Thr Glu Lys Trp Gln Asp Ile Ile Lys Glu Val Lys Phe Leu
65 70 75 80

Gln Arg Ile Lys His Pro Asn Ser Ile Glu Tyr Lys Gly Cys Tyr Leu
85 90 95

Arg Glu His Thr Ala Trp Leu Val Met Glu Tyr Cys Leu Gly Ser Ala
100 105 110

Ser Asp Leu Leu Glu Val His Lys Lys Pro Leu Gln Glu Val Glu Ile
115 120 125

Ala Ala Ile Thr His Gly Ala Leu Gln Gly Leu Ala Tyr Leu His Ser
130 135 140

His Thr Met Ile His Arg Asp Ile Lys Ala Gly Asn Ile Leu Leu Thr
145 150 155 160

Glu Pro Gly Gln Val Lys Leu Ala Asp Phe Gly Ser Ala Ser Met Ala
165 170 175

Ser Pro Ala Asn Ser Phe Val Gly Thr Pro Tyr Trp Met Ala Pro Glu
180 185 190

Val Ile Leu Ala Met Asp Glu Gly Gln Tyr Asp Gly Lys Val Asp Val
195 200 205

Trp Ser Leu Gly Ile Thr Cys Ile Glu Leu Ala Glu Arg Lys Pro Pro

60

210	215	220
Leu Phe Asn Met Asn Ala Met Ser Ala Leu Tyr His Ile Ala Gln Asn 225	230	235
Glu Ser Pro Thr Leu Gln Ser Asn Glu Trp Ser Asp Tyr Phe Arg Asn 245	250	255
Phe Val Asp Ser Cys Leu Gln Lys Ile Pro Gln Asp Arg Pro Thr Ser 260	265	270
Glu Glu Leu Leu Lys His Ile Phe Val Leu Arg Glu Arg Pro Glu Thr 275	280	285
Val Leu Ile Asp Leu Ile Gln Arg Thr Lys Asp Ala Val Arg Glu Leu 290	295	300
Asp Asn Leu Gln Tyr Arg Lys Met Lys Lys Leu Leu Phe Gln Glu Ala 305	310	315
His Asn Gly Pro Ala Val Glu Ala Gln Glu Glu Glu Glu Gln Asp 325	330	335
His Gly Val Gly Arg Thr Gly Thr Val Asn Ser Val Gly Ser Asn Gln 340	345	350
Ser Ile Pro Ser Met Ser Ile Ser Ala Ser Ser Gln Ser Ser Val 355	360	365
Asn Ser Leu Pro Asp Val Ser Asp Asp Lys Ser Glu Leu Asp Met Met 370	375	380
Glu Gly Asp His Thr Val Met Ser Asn Ser Ser Val Ile His Leu Lys 385	390	395
Pro Glu Glu Glu Asn Tyr Arg Glu Glu Gly Asp Pro Arg Thr Arg Ala 405	410	415
Ser Asp Pro Gln Ser Pro Pro Gln Val Ser Arg His Lys Ser His Tyr 420	425	430
Arg Asn Arg Glu His Phe Ala Thr Ile Arg Thr Ala Ser Leu Val Thr 435	440	445
Arg Gln Met Gln Glu His Glu Gln Asp Ser Glu Leu Arg Glu Gln Met 450	455	460
Ser Gly Tyr Lys Arg Met Arg Arg Gln His Gln Lys Gln Leu Met Thr 465	470	475
Leu Glu Asn Lys Leu Lys Ala Glu Met Asp Glu His Arg Leu Arg Leu 485	490	495
Asp Lys Asp Leu Glu Thr Gln Arg Asn Asn Phe Ala Ala Glu Met Glu 500	505	510

Lys Leu Ile Lys Lys His Gln Ala Ala Met Glu Lys Glu Ala Lys Val
 515 520 525

Met Ser Asn Glu Glu Lys Lys Phe Gln Gln His Ile Gln Ala Gln Gln
 530 535 540

Lys Lys Glu Leu Asn Ser Phe Leu Glu Ser Gln Lys Arg Glu Tyr Lys
 545 550 555 560

Leu Arg Lys Glu Gln Leu Lys Glu Glu Leu Asn Glu Asn Gln Ser Thr
 565 570 575

Pro Lys Lys Glu Lys Gln Glu Trp Leu Ser Lys Gln Lys Glu Asn Ile
 580 585 590

Gln His Phe Gln Ala Glu Glu Ala Asn Leu Leu Arg Arg Gln Arg
 595 600 605

Gln Tyr Leu Glu Leu Glu Cys Arg Arg Phe Lys Arg Arg Met Leu Leu
 610 615 620

Gly Arg His Asn Leu Glu Gln Asp Leu Val Arg Glu Glu Leu Asn Lys
 625 630 635 640

Arg Gln Thr Gln Lys Asp Leu Glu His Ala Met Leu Leu Arg Gln His
 645 650 655

Glu Ser Met Gln Glu Leu Glu Phe Arg His Leu Asn Thr Ile Gln Lys
 660 665 670

Met Arg Cys Glu Leu Ile Arg Leu Gln His Gln Thr Glu Leu Thr Asn
 675 680 685

Gln Leu Glu Tyr Asn Lys Arg Arg Glu Arg Glu Leu Arg Arg Lys His
 690 695 700

Val Met Glu Val Arg Gln Gln Pro Lys Ser Leu Lys Ser Lys Glu Leu
 705 710 715 720

Gln Ile Lys Lys Gln Phe Gln Asp Thr Cys Lys Ile Gln Thr Arg Gln
 725 730 735

Tyr Lys Ala Leu Arg Asn His Leu Leu Glu Thr Thr Pro Lys Ser Glu
 740 745 750

His Lys Ala Val Leu Lys Arg Leu Lys Glu Glu Gln Thr Arg Lys Leu
 755 760 765

Ala Ile Leu Ala Glu Gln Tyr Asp His Ser Ile Asn Glu Met Leu Ser
 770 775 780

Thr Gln Ala Leu Arg Leu Asp Glu Ala Gln Glu Ala Glu Cys Gln Val
 785 790 795 800

Leu Lys Met Gln Leu Gln Gln Glu Leu Glu Leu Asn Ala Tyr Gln

62

805	810	815
Ser Lys Ile Lys Met Gln Ala Glu Ala Gln His Asp Arg Glu Leu Arg		
820	825	830
Glu Leu Glu Gln Arg Val Ser Leu Arg Arg Ala Leu Leu Glu Gln Lys		
835	840	845
Ile Glu Glu Glu Met Leu Ala Leu Gln Asn Glu Arg Thr Glu Arg Ile		
850	855	860
Arg Ser Leu Leu Glu Arg Gln Ala Arg Glu Ile Glu Ala Phe Asp Ser		
865	870	875
Glu Ser Met Arg Leu Gly Phe Ser Asn Met Val Leu Ser Asn Leu Ser		
885	890	895
Pro Glu Ala Phe Ser His Ser Tyr Pro Gly Ala Ser Gly Trp Ser His		
900	905	910
Asn Pro Thr Gly Gly Pro Gly Pro His Trp Gly His Pro Met Gly Gly		
915	920	925
Pro Pro Gln Ala Trp Gly His Pro Met Gln Gly Gly Pro Gln Pro Trp		
930	935	940
Gly His Pro Ser Gly Pro Met Gln Gly Val Pro Arg Gly Ser Ser Met		
945	950	955
Gly Val Arg Asn Ser Pro Gln Ala Leu Arg Arg Thr Ala Ser Gly Gly		
965	970	975
Arg Thr Glu Gln Gly Met Ser Arg Ser Thr Ser Val Thr Ser Gln Ile		
980	985	990
Ser Asn Gly Ser His Met Ser Tyr Thr		
995	1000	

<210> 32

<211> 25

<212> DNA

<213> TRK1

<220>

<223> "n" stands for a, c, g or t.

"y" stands for c or t.

"r" stands for a or g.

<400> 32

ctgaattcgg ngcnttyggm aargt

25

<210> 33
<211> 24
<212> DNA
<213> TRK4

<220>
<223> "n" stands for a, c, g or t.
"y" stands for c or t.

<400> 33

gctggatccy tcnggnggca tcca

24

<210> 34
<211> 23
<212> DNA
<213> ROS1

<220>
<223> "n" stands for a, c, g or t.
"y" stands for c or t.
"r" stands for a or g.

<400> 34

gcnttyggng argtntayga rgg

23

<210> 35
<211> 24
<212> DNA
<213> CCK4b

<220>
<223> "n" stands for a, c, g or t.
"y" stands for c or t.
"s" stands for c or g.
"w" stands for a or t.

<400> 35

gctggatccy tcnggnswca tcca

24

<210> 36
<211> 23
<212> DNA
<213> CCK4c

64

<220>
<223> "n" stands for a, c, g or t.
"y" stands for c or t.
"r" stands for a or g.

<400> 36

gagtttyggng argtnttyyt ngc

23

<210> 37
<211> 6
<212> PRT
<213> TRK1

<400> 37

Gly Ala Phe Gly Lys Val
1 5

<210> 38
<211> 5
<212> PRT
<213> TRK4

<400> 38

Trp Met Pro Pro Glu
1 5

<210> 39
<211> 8
<212> PRT
<213> ROS1

<400> 39

Ala Phe Gly Glu Val Tyr Glu Gly
1 5

<210> 40
<211> 5
<212> PRT
<213> CCK4b

<400> 40

Trp Met Ser Pro Glu
1 5

<210> 41
<211> 8
<212> PRT
<213> CCK4c

<400> 41

Glu Phe Gly Glu Val Tyr Glu Gly
1 5

<210> 42
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 42

cacagaaaacg gtcagattca c

21

<210> 43
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 43

gatcagggtg acatcaaggg ac

22

<210> 44
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 44

ctcatctgta cacacttcat gg

22

<210> 45

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 45

gattccacaca ctgttagatgt c

21

<210> 46

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 46

ggccctcgac tacatccacc acat

24

<210> 47

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 47

caacgaaaact aacacagcat aagg

24

<210> 48

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

67

<223> Synthesized nucleic acid molecule

<400> 48

atggcgaacg actctccgc gaa

23

<210> 49

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 49

acaccaaaaat caacaagttt cacctc

26

<210> 50

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 50

agttaacaagg aattccaagt tct

23

<210> 51

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 51

atgaagagga agaaatcaa ctg

23

<210> 52

<211> 20

68

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 52

agatggactg tactgggagg

20

<210> 53

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 53

acttttgca gctctgtggg

20

<210> 54

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 54

aaggttatgg atgtcacagg g

21

<210> 55

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 55

ctcacaaggt tgccaacagg

20

69

<210> 56
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 56

agtccccacc agaaggttta c

21

<210> 57
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 57

tcaggggtca gaggtcacg

19

<210> 58
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 58

cccaaaccct accacaaaatt c

21

<210> 59
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 59

cccccgaa acgatgacca

20

70

<210> 60
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 60

agccgctgcc cctcctctac tgt

23

<210> 61
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 61

accgcaacat cgccacctac tac

23

<210> 62
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 62

ctcgacgtcg tggaccacc

19

<210> 63
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 63

caatgttaac ccactctatg tctc

24

<210> 64
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 64

agtttgcga tgggtttctt ttc

23

<210> 65
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 65

ccgccccatgaa ccccccggctt

19

<210> 66
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 66

cgattgccaa agaccgtgtc a

21

<210> 67
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 67

agaagttgca gctgttgaga gga

23

<210> 68
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 68

tatggccgt gtaaggattt c 21

<210> 69
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 69

gtgccagaag ttttgttggaa 23

<210> 70
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 70

tattgaattg gcggAACCGGA aq 22

<210> 71
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 71

ttgttttqfq ctcattcttt qqaq

<210> 72
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 72

gattgctttg tgctcattct ttgg

24

<210> 73
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 73

ttgttctaag agtgccctcc g

21

<210> 74
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 74

aagaccatgc cgtgcgccg

19

<210> 75
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 75

atcccttcag gttctggta tgg

23

<210> 76
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule
<400> 76

agtccgccaa gcctcaatg

19

<210> 77
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule
<400> 77

caccttcagc tgctctatca t

21

<210> 78
<211> 12
<212> PRT
<213> Peptide 540A

<400> 78

His Gly Asp Pro Arg Pro Glu Pro Arg Pro Thr Gln
1 5 10

<210> 79
<211> 10
<212> PRT
<213> Peptide 539A

<400> 79

Cys Leu Asp Phe Pro Lys Glu Asp Tyr Arg
1 5 10

<210> 80
<211> 19
<212> PRT
<213> Peptide 541A

<400> 80

-Asp Pro Arg Thr Arg Ala Ser Asp Pro Gln Ser Pro Pro Gln Val Ser
1 5 10 15

Arg His Lys

<210> 81
<211> 14
<212> PRT
<213> Peptide 542A

<400> 81

Pro Ser Thr Asn Arg Ala Gly Ser Leu Lys Asp Pro Glu Cys
1 5 10

<210> 82
<211> 15
<212> PRT
<213> Peptide 554A

<400> 82

Cys Leu Val Pro Leu Ile Gln Leu Tyr Arg Lys Gin Thr Ser Thr
1 5 10 15

<210> 83
<211> 10
<212> PRT
<213> Peptide 579A

<400> 83

Cys Pro Leu Met Arg Gln Asn Arg Thr Arg
1 5 10

<210> 84
<211> 426

<212> PRT
<213> STE20h

<400> 84

Met Ala His Leu Arg Gly Phe Ala Asn Gln His Ser Arg Val Asp Pro
1 5 10 15

Glu Glu Leu Phe Thr Lys Leu Asp Arg Ile Gly Lys Gly Ser Phe Gly
20 25 30

Glu Val Tyr Lys Gly Ile Asp Asn His Thr Lys Glu Val Val Ala Ile
35 40 45

Lys Ile Ile Asp Leu Glu Glu Ala Glu Asp Glu Ile Glu Asp Ile Gln
50 55 60

Gln Glu Ile Thr Val Leu Ser Gln Cys Asp Ser Pro Tyr Ile Thr Arg
65 70 75 80

Tyr Phe Gly Ser Tyr Leu Lys Ser Thr Lys Leu Trp Ile Ile Met Glu
85 90 95

Tyr Leu Gly Gly Ser Ala Leu Asp Leu Leu Lys Pro Gly Pro Leu
100 105 110

Glu Glu Thr Tyr Ile Ala Thr Ile Leu Arg Glu Ile Leu Lys Gly Leu
115 120 125

Asp Tyr Leu His Ser Glu Arg Lys Ile His Arg Asp Ile Lys Ala Ala
130 135 140

Asn Val Leu Leu Ser Glu Gln Gly Asp Val Lys Leu Ala Asp Phe Gly
145 150 155 160

Val Ala Gly Gln Leu Thr Asp Thr Gln Ile Lys Arg Asn Thr Phe Val
165 170 175

Gly Thr Pro Phe Trp Met Ala Pro Glu Val Ile Lys Gln Ser Ala Tyr
180 185 190

Asp Phe Lys Ala Asp Ile Trp Ser Leu Gly Ile Thr Ala Ile Glu Leu
195 200 205

Ala Lys Gly Glu Pro Pro Asn Ser Asp Leu His Pro Met Arg Val Leu
210 215 220

Phe Leu Ile Pro Lys Asn Ser Pro Pro Thr Leu Glu Gly Gln His Ser
225 230 235 240

Lys Pro Phe Lys Glu Phe Val Glu Ala Cys Leu Asn Lys Asp Pro Arg
245 250 255

Phe Arg Pro Thr Ala Lys Glu Leu Leu Lys His Lys Phe Ile Thr Arg
260 265 270

Tyr Thr Lys Lys Thr Ser Phe Leu Thr Glu Leu Ile Asp Arg Tyr Lys
275 280 285

Arg Trp Lys Ser Glu Gly His Gly Glu Glu Ser Ser Ser Glu Asp Ser
290 295 300

Asp Ile Asp Gly Glu Ala Glu Asp Gly Glu Gln Gly Pro Ile Trp Thr
305 310 315 320

Phe Pro Pro Thr Ile Arg Pro Ser Pro His Ser Lys Leu His Lys Gly
325 330 335

Thr Ala Leu His Ser Ser Gln Lys Pro Ala Glu Pro Val Lys Arg Gln
340 345 350

Pro Arg Ser Gln Cys Leu Ser Thr Leu Val Arg Pro Val Phe Gly Glu
355 360 365

Leu Lys Glu Lys His Lys Gln Ser Gly Gly Ser Val Gly Ala Leu Glu
370 375 380

Glu Leu Glu Asn Ala Phe Ser Leu Ala Glu Glu Ser Cys Pro Gly Ile
385 390 395 400

Ser Asp Lys Leu Met Val His Leu Val Glu Arg Val Gln Arg Phe Ser
405 410 415

His Asn Arg Asn His Leu Thr Ser Thr Arg
420 425

<210> 85
<211> 431
<212> PRT
<213> MST3

<400> 85

Met Ala His Ser Pro Val Gln Ser Gly Leu Pro Gly Met Gln Asn Leu
1 5 10 15

Lys Ala Asp Pro Glu Glu Leu Phe Thr Lys Leu Glu Lys Ile Gly Lys
20 25 30

Gly Ser Phe Gly Glu Val Phe Lys Gly Ile Asp Asn Arg Thr Gln Lys
35 40 45

Val Val Ala Ile Lys Ile Ile Asp Leu Glu Glu Ala Glu Asp Glu Ile
50 55 60

Glu Asp Ile Gln Gln Glu Ile Thr Val Leu Ser Gln Cys Asp Ser Pro
65 70 75 80

Tyr Val Thr Lys Tyr Tyr Gly Ser Tyr Leu Lys Asp Thr Lys Leu Trp

78

85

90

95

Ile Ile Met Glu Tyr Leu Gly Gly Ser Ala Leu Asp Leu Leu Glu
 100 105 110

Pro Gly Pro Leu Asp Glu Thr Gln Ile Ala Thr Ile Leu Arg Glu Ile
 115 120 125

Leu Lys Gly Leu Asp Tyr Leu His Ser Glu Lys Lys Ile His Arg Asp
 130 135 140

Ile Lys Ala Ala Asn Val Leu Leu Ser Glu His Gly Glu Val Lys Leu
 145 150 155 160

Ala Asp Phe Gly Val Ala Gly Gln Leu Thr Asp Thr Gln Ile Lys Arg
 165 170 175

Asn Thr Phe Val Gly Thr Pro Phe Trp Met Ala Pro Glu Val Ile Lys
 180 185 190

Gln Ser Ala Tyr Asp Ser Lys Ala Asp Ile Trp Ser Leu Gly Ile Thr
 195 200 205

Ala Ile Glu Leu Ala Arg Gly Glu Pro Pro His Ser Glu Leu His Pro
 210 215 220

Met Lys Val Leu Phe Leu Ile Pro Lys Asn Asn Pro Pro Thr Leu Glu
 225 230 235 240

Gly Asn Tyr Ser Lys Pro Leu Lys Glu Phe Val Glu Ala Cys Leu Asn
 245 250 255

Lys Glu Pro Ser Phe Arg Pro Thr Ala Lys Glu Leu Leu Lys His Lys
 260 265 270

Phe Ile Leu Arg Asn Ala Lys Lys Thr Ser Tyr Leu Thr Glu Leu Ile
 275 280 285

Asp Arg Tyr Lys Arg Trp Lys Ala Glu Gln Ser His Asp Asp Ser Ser
 290 295 300

Ser Glu Asp Ser Asp Ala Glu Thr Asp Gly Gln Ala Ser Gly Gly Ser
 305 310 315 320

Asp Ser Gly Asp Trp Ile Phe Thr Ile Arg Glu Lys Asp Pro Lys Asn
 325 330 335

Leu Glu Asn Gly Ala Leu Gln Pro Ser Asp Leu Asp Arg Asn Lys Met
 340 345 350

Lys Asp Ile Pro Lys Arg Pro Phe Ser Gln Cys Leu Ser Thr Ile Ile
 355 360 365

Ser Pro Leu Phe Ala Glu Leu Lys Glu Lys Ser Gln Ala Cys Gly Gly
 370 375 380

Asn Leu Gly Ser Ile Glu Glu Leu Arg Gly Ala Ile Tyr Leu Ala Glu
 385 390 395 400

Glu Ala Cys Pro Gly Ile Ser Asp Thr Met Val Ala Gln Leu Val Gln
 405 410 415

Arg Leu Gln Arg Tyr Ser Leu Ser Gly Gly Gly Thr Ser Ser His
 420 425 430

<210> 86
 <211> 443
 <212> PRT
 <213> T19A5.2_ce

<220>

<223> "Xaa" stands for any amino acid.

<400> 86

Met Thr Thr Thr Ser Ser Asp Glu Leu Pro Arg Gln Ala Asp Asp Asp
 1 5 10 15

Ser Met Lys Trp Asp Arg Ile Tyr Ile Gln Lys Leu Asp Pro Glu Val
 20 25 30

Ile Phe Thr Lys Gln Glu Arg Ile Gly Arg Gly Ser Phe Gly Glu Val
 35 40 45

Tyr Lys Gly Ile Asp Asn Arg Thr Gly Arg Val Val Ala Ile Lys Ile
 50 55 60

Ile Asp Leu Glu Gln Ala Glu Asp Glu Ile Glu Asp Ile Gln Gln Glu
 65 70 75 80

Ile Gln Val Leu Ser Gln Cys Asp Ser Gln Tyr Val Thr Lys Tyr Phe
 85 90 95

Gly Ser Phe Leu Lys Gly Ser Lys Leu Trp Ile Ile Met Glu Tyr Leu
 100 105 110

Gly Gly Gly Ser Ala Leu Asp Leu Thr Lys Ser Gly Lys Leu Asp Glu
 115 120 125

Ser His Ile Ala Val Ile Leu Arg Glu Ile Leu Lys Gly Leu Glu Tyr
 130 135 140

Leu His Ser Glu Arg Lys Ile His Arg Asp Ile Lys Ala Ala Asn Val
 145 150 155 160

Leu Val Ser Glu His Gly Asp Val Lys Val Ala Asp Phe Gly Val Ala
 165 170 175

80

Gly Gln Leu Thr Glu Thr Val Lys Lys Arg Ile Thr Phe Val Gly Ser
180 185 190

Pro Phe Trp Met Ala Pro Glu Leu Ile Lys Gln Ser Ser Tyr Asp Tyr
195 200 205

Lys Ala Asp Ile Trp Ser Leu Gly Ile Thr Ala Ile Glu Leu Ala Asn
210 215 220

Gly Glu Pro Pro His Ser Asp Leu His Pro Met Arg Val Leu Phe Leu
225 230 235 240

Ile Pro Lys Asn Pro Pro Val Leu Gln Gly Ser Gln Trp Ser Lys
245 250 255

Pro Phe Lys Glu Phe Val Glu Met Cys Leu Asn Lys Asp Pro Glu Asn
260 265 270

Arg Pro Ser Ala Ser Thr Leu Leu Lys His Gln Phe Ile Lys Arg Ala
275 280 285

Lys Lys Asn Ser Ile Leu Val Asp Leu Ile Glu Arg Ala Ala Glu Tyr
290 295 300

Arg Leu Arg Thr Gly Val Ser Ser Asp Ser Asp Leu Asp Glu Asp Ser
305 310 315 320

Asp Gly Gly Gly Thr Ser Lys Trp Asp Tyr Pro Thr Val Arg Gly
325 330 335

Pro Arg Val Ser Ala Asp Asp Asp Gly Thr Val Arg Gln Arg Thr Asp
340 345 350

Arg Pro Arg Ala Gln Val Asp Arg Arg Ser Pro Ser Gly Ser Pro Gly
355 360 365

Gly Thr Ile Val Arg Gly Ser Pro Gln Val Ala Ala Val Ala Glu Gln
370 375 380

Leu Arg Asn Ser Xaa Xaa Ala Leu Asp Gln Leu Arg His Val Phe Arg
385 390 395 400

Asp Val Glu Asp Ser Cys Pro Gly Ile Cys Asn Glu Leu Ile Glu Glu
405 410 415

Leu Met Gln Arg Ile Ala Val Pro Gln Val Ser Gln Ser Asp Leu Asp
420 425 430

Ala Ala Ile Arg Arg Leu Thr Thr Pro Pro Ser
435 440

<211> 275
<212> PRT
<213> Pak_sp

<400> 87

Leu Leu Tyr Arg Asn Phe Val Lys Ile Gly Gln Gly Ala Ser Gly Asp
1 5 10 15

Val Tyr Ser Ala Arg Gln Val Gly Thr Asn Leu Ser Val Ala Ile Lys
20 25 30

Lys Met Asn Ile Asn Gln Gln Pro Lys Lys Glu Phe Ile Val Asn Glu
35 40 45

Ile Leu Val Met Lys Ser His His Lys Asn Ile Val Asn Phe Ile
50 55 60

Asp Thr Phe Phe Tyr Lys Ser Glu Leu Trp Met Val Met Glu Tyr Met
65 70 75 80

Arg Gly Gly Ser Leu Thr Glu Val Val Thr Asn Asn Thr Leu Ser Glu
85 90 95

Gly Gln Ile Ala Ala Ile Cys Lys Glu Thr Leu Glu Gly Leu Gln His
100 105 110

Leu His Glu Asn Gly Ile Val His Arg Asp Ile Lys Ser Asp Asn Ile
115 120 125

Leu Leu Ser Leu Gln Gly Asp Ile Lys Leu Thr Asp Phe Gly Phe Cys
130 135 140

Ala Gln Ile Asp Ser Asn Met Thr Lys Arg Thr Thr Met Val Gly Thr
145 150 155 160

Pro Tyr Trp Met Ala Pro Glu Val Val Thr Arg Lys Glu Tyr Gly Phe
165 170 175

Lys Val Asp Val Trp Ser Leu Gly Ile Met Ala Ile Glu Met Val Glu
180 185 190

Gly Glu Pro Pro Tyr Leu Asn Glu Asn Pro Leu Arg Ala Leu Tyr Leu
195 200 205

Ile Ala Thr Ile Gly Thr Pro Lys Ile Ser Arg Pro Glu Leu Leu Ser
210 215 220

Ser Val Phe His Asp Phe Leu Ser Lys Ser Leu Thr Val Asn Pro Lys
225 230 235 240

Gln Arg Pro Ser Ser Gly Glu Leu Leu Arg His Pro Phe Leu Lys Gln
245 250 255

Ala Val Pro Val Ser Ser Leu Ile Pro Leu Ile Lys Ser Ile His His
260 265 270

Ser Gly Lys
275

<210> 88
<211> 1109
<212> PRT
<213> ZC504.4_ce

<400> 88

Met Ser Ser Ser Gly Leu Asp Glu Ile Asp Leu Asn Ser Leu Arg Asp
1 5 10 15

Pro Ala Gly Ile Phe Glu Leu Ile Glu Val Val Gly Asn Gly Thr Tyr
20 25 30

Gly Gln Val Tyr Lys Gly Arg His Val Lys Thr Ala Gln Leu Ala Ala
35 40 45

Ile Lys Ile Met Asn Ile Asn Glu Asp Glu Asp Glu Ile Lys Leu
50 55 60

Glu Ile Asn Met Leu Lys Lys His Ser His His Arg Asn Val Ala Thr
65 70 75 80

Tyr Tyr Gly Ala Phe Ile Lys Lys Leu Pro Ser Ser Thr Gly Lys His
85 90 95

Asp Gln Leu Trp Leu Val Met Glu Phe Cys Gly Ser Gly Ser Ile Thr
100 105 110

Asp Leu Val Lys Asn Thr Lys Gly Gly Ser Leu Lys Glu Glu Trp Ile
115 120 125

Ala Tyr Ile Cys Arg Glu Ile Leu Arg Gly Leu Tyr His Leu His Gln
130 135 140

Ser Lys Val Ile His Arg Asp Ile Lys Gly Gin Asn Val Leu Leu Thr
145 150 155 160

Asp Ser Ala Glu Val Lys Leu Val Asp Phe Gly Val Ser Ala Gln Leu
165 170 175

Asp Lys Thr Val Gly Arg Arg Asn Thr Phe Ile Gly Thr Pro Tyr Trp
180 185 190

Met Ala Pro Glu Val Ile Ala Cys Asp Glu Ser Pro Glu Ala Thr Tyr
195 200 205

Asp Ser Arg Ser Asp Leu Trp Ser Leu Gly Ile Thr Ala Leu Glu Met
210 215 220

Ala Glu Gly His Pro Pro Leu Cys Asp Met His Pro Met Arg Ala Leu
 225 230 235 240

Phe Leu Ile Pro Arg Asn Pro Pro Pro Lys Leu Lys Arg Asn Lys Lys
 245 250 255

Trp Thr Lys Lys Phe Glu Thr Phe Ile Glu Thr Val Leu Val Lys Asp
 260 265 270

Tyr His Gln Arg Pro Tyr Thr Gly Ala Leu Leu Arg His Pro Phe Ile
 275 280 285

Lys Glu Gln Pro His Glu Gln Thr Ile Arg His Ser Ile Lys Glu His
 290 295 300

Ile Asp Arg Asn Arg Arg Val Lys Lys Asp Asp Ala Asp Tyr Glu Tyr
 305 310 315 320

Ser Gly Ser Glu Asp Asp Glu Pro Ser Pro Asn Asn Arg Asp Asp Ser
 325 330 335

Glu Ser Ser Ser Met Ile Pro Met Asp Asn Thr Leu Arg Lys Gly Phe
 340 345 350

Gln Lys Leu Gln Glu Ser Ser Arg Gly Phe Ala Glu Pro Gly Ala Gln
 355 360 365

Gln Leu Arg Arg Leu Pro Gln Gln Pro Ala Pro Ala Pro Phe Gln Tyr
 370 375 380

Gln Gln Ser Arg Tyr Val Glu Pro Arg Arg Glu Ser Ser Glu Val Lys
 385 390 395 400

Leu Arg Ala Val Ser Ser Arg Gly Ala Ala Asp Gly Pro Arg His Ser
 405 410 415

Pro Ala Ser Arg Pro Arg Pro Arg Ser Pro Gln Gln Ser His Pro Ala
 420 425 430

Ala Pro His Leu Ala Asp Leu Ala Asn Tyr Glu Lys Arg Arg Arg Ser
 435 440 445

Glu Arg Glu Glu Arg Arg Glu Arg Glu Arg Gln Ala His His Ala Met
 450 455 460

Pro Ile Ala Arg Val Ser Ala Ser Val Pro Ala Pro Gln Gln Ser Arg
 465 470 475 480

Lys Met Ser Glu Pro Leu Leu Ile Thr His Val Lys Pro Glu Asp Leu
 485 490 495

Asp Val Leu Ala Ser Glu Leu Ser Lys Met Gly Gly His His Asn Gly
 500 505 510

Arg Ser Arg Glu Glu Ser Met Ser Pro Pro Pro Pro Ala Pro Pro Pro

515	520	525
Arg Glu Ala Ser Ile Ser Ser Ile Thr Asp Thr Ile Asp Val Gly Glu		
530	535	540
Leu Asp Asn Gly Ala Asp Ala Glu Trp Asp Asp Leu Lys Asp Ile Met		
545	550	555
560		
Met Asn Gly Glu Gly Thr Leu Arg Gly Pro Asn Lys Pro Leu Pro Pro		
565	570	575
Thr Pro Thr Asp Gly Glu Asn Thr Leu Val Ser Asp Val Arg Arg Asn		
580	585	590
Gly Asn Gly Asn Ser Gly His Gly Ala Tyr Lys Gly Lys Lys Ile Pro		
595	600	605
Glu Ile Arg Pro Gly Ile Ile Ser Leu Asp Asp Asp Ser Asp Ser		
610	615	620
Asp Asn Glu Glu Gly Asn Glu Pro Leu Met Phe Lys Pro Ile Val Arg		
625	630	635
640		
Cys Pro Phe Ser Ile Phe Phe Trp Phe Leu Ser Ala Asn Val Ile His		
645	650	655
Ser Val Asp Gly Ser Ile Pro Leu Val Lys His Leu Ile Trp Phe Gln		
660	665	670
Asn Ala Ser Ser Ser Arg Gly Ala Leu Pro Asp Leu Leu Pro Lys Ser		
675	680	685
Pro Asp Leu Arg Arg Gln Ile Asn Asp Gln Thr Arg Gln Met Ser Asp		
690	695	700
Asp Arg Ala Asp Glu Gln Pro Asn Gly Phe Gln Asn Ser Asp Ser Arg		
705	710	715
720		
Ser Ser Ile Gln His Ser Phe Ser Asn Arg Asp Arg Glu Lys Ser Phe		
725	730	735
Val Gly Tyr Phe Gly Gly Ala Gly Ala Gly Gly Thr Val Asn		
740	745	750
Arg Pro Gly Arg Pro Gln Asp Ile Asn Gln Val Gln Val Asn Val Thr		
755	760	765
Pro Asn Ser Asn Gly Thr Pro Ala Glu Asn Asp Ala Pro Glu Ile Arg		
770	775	780
Lys Tyr Lys Lys Phe Ser Gly Glu Ile Leu Cys Ala Ala Leu Trp		
785	790	795
800		
Gly Val Asn Leu Leu Ile Gly Thr Asp Ser Gly Leu Met Leu Leu Asp		
805	810	815

Arg Ser Gly Gln Gly Lys Val Tyr Pro Leu Ile Ser Arg Arg Arg Phe
 820 825 830
 Asp Gln Met Thr Val Leu Glu Gly Gln Asn Ile Leu Ala Thr Ile Ser
 835 840 845
 Gly Arg Lys Arg Arg Ile Arg Val Tyr Tyr Leu Ser Trp Leu Arg Gln
 850 855 860
 Lys Ile Leu Arg Thr Glu Gly Ala Gly Ser Ala Asn Thr Thr Glu Lys
 865 870 875 880
 Arg Asn Gly Trp Val Asn Val Gly Asp Leu Gln Gly Ala Ile His Phe
 885 890 895
 Lys Ile Val Arg Tyr Glu Arg Ile Lys Phe Leu Val Val Gly Leu Glu
 900 905 910
 Ser Ser Ile Glu Ile Tyr Ala Trp Ala Pro Lys Pro Tyr His Lys Phe
 915 920 925
 Met Ser Phe Lys Ser Phe Gly Ser Leu Ser His Val Pro Leu Ile Val
 930 935 940
 Asp Leu Thr Val Glu Asp Asn Ala Arg Leu Lys Val Leu Tyr Gly Ser
 945 950 955 960
 Thr Gly Gly Phe His Ala Ile Asp Leu Asp Ser Ala Ala Val Tyr Asp
 965 970 975
 Ile Tyr Thr Pro Ala Gln Ser Gly Gln Thr Thr Thr Pro His Cys Ile
 980 985 990
 Val Val Leu Pro Asn Ser Asn Gly Met Gln Leu Leu Leu Cys Tyr Asp
 995 1000 1005
 Asn Glu Gly Val Tyr Val Asn Thr Tyr Gly Arg Met Thr Lys Asn Val
 1010 1015 1020
 Val Leu Gln Trp Gly Glu Met Pro Ser Ser Val Ala Tyr Ile Ser Thr
 1025 1030 1035 1040
 Gly Gln Ile Met Gly Trp Gly Asn Lys Ala Ile Glu Ile Arg Ser Val
 1045 1050 1055
 Asp Thr Gly His Leu Asp Gly Val Phe Met His Lys Lys Ala Gln Lys
 1060 1065 1070
 Leu Lys Phe Leu Cys Glu Arg Asn Asp Lys Val Phe Phe Ser Ser Ala
 1075 1080 1085
 Lys Gly Gly Ser Cys Gln Ile Tyr Phe Met Thr Leu Asn Lys Pro
 1090 1095 1100

Gly Leu Thr Asn Trp
1105

<210> 89
<211> 1233
<212> PRT
<213> NIK_m

<400> 89

Met Ala Asn Asp Ser Pro Ala Lys Ser Leu Val Asp Ile Asp Leu Ser
1 5 10 15

Ser Leu Arg Asp Pro Ala Gly Ile Phe Glu Leu Val Glu Val Val Gly
20 25 30

Asn Gly Thr Tyr Gly Gln Val Tyr Lys Gly Arg His Val Lys Thr Val
35 40 45

Thr Ala Ala Ile Lys Val Met Asp Val Thr Glu Asp Glu Glu Glu
50 55 60

Ile Thr Leu Glu Ile Asn Met Leu Lys Tyr Ser His His Arg Asn
65 70 75 80

Ile Ala Thr Tyr Tyr Gly Ala Phe Ile Lys Lys Ser Pro Pro Gly His
85 90 95

Asp Asp Gln Leu Trp Leu Val Met Glu Phe Cys Gly Ala Gly Ser Ile
100 105 110

Thr Asp Leu Val Lys Asn Thr Lys Gly Asn Thr Leu Lys Glu Asp Trp
115 120 125

Ile Ala Tyr Ile Ser Arg Glu Ile Leu Arg Gly Leu Ala His Leu His
130 135 140

Ile His His Val Ile His Arg Asp Ile Lys Gly Gln Asn Val Leu Leu
145 150 155 160

Thr Glu Asn Ala Glu Val Lys Leu Val Asp Phe Gly Val Ser Ala Gln
165 170 175

Leu Asp Arg Thr Val Gly Arg Arg Asn Thr Phe Ile Gly Thr Pro Tyr
180 185 190

Trp Met Ala Pro Glu Val Ile Ala Cys Asp Glu Asn Pro Asp Ala Thr
195 200 205

Tyr Asp Tyr Arg Ser Asp Leu Trp Ser Cys Gly Ile Thr Ala Ile Glu
210 215 220

Met Ala Glu Gly Gly Pro Pro Leu Cys Asp Met His Pro Met Arg Ala

225	230	235	240
Leu Phe Leu Ile Pro Arg Asn Pro Pro Pro Arg Leu Lys Ser Lys Lys			
245		250	255
Trp Ser Lys Lys Phe Phe Ser Phe Ile Glu Gly Cys Leu Val Lys Asn			
260		265	270
Tyr Met Gln Arg Pro Ser Thr Glu Gln Leu Leu Lys His Pro Phe Ile			
275	280	285	
Arg Asp Gln Pro Asn Glu Arg Gln Val Arg Ile Gln Leu Lys Asp His			
290	295	300	
Ile Asp Arg Thr Arg Lys Lys Arg Gly Glu Lys Asp Glu Thr Glu Tyr			
305	310	315	320
Glu Tyr Ser Gly Ser Glu Glu Glu Glu Glu Val Pro Glu Gln Glu			
325		330	335
Gly Glu Pro Ser Ser Ile Val Asn Val Pro Gly Glu Ser Thr Leu Arg			
340	345	350	
Arg Asp Phe Leu Arg Leu Gln Gln Glu Asn Lys Glu Arg Ser Glu Ala			
355	360	365	
Leu Arg Arg Gln Gln Leu Leu Gln Glu Gln Gln Leu Arg Glu Gln Glu			
370	375	380	
Glu Tyr Lys Arg Gln Leu Leu Ala Glu Arg Gln Lys Arg Ile Glu Gln			
385	390	395	400
Gln Lys Glu Gln Arg Arg Leu Glu Glu Gln Gln Arg Arg Glu Arg			
405		410	415
Glu Ala Arg Arg Gln Gln Glu Arg Glu Gln Arg Arg Arg Glu Gln Glu			
420	425	430	
Glu Lys Arg Arg Leu Glu Glu Leu Glu Arg Arg Arg Lys Glu Glu Glu			
435	440	445	
Glu Arg Arg Arg Ala Glu Glu Glu Lys Arg Arg Val Glu Arg Glu Gln			
450	455	460	
Glu Tyr Ile Arg Arg Gln Leu Glu Glu Glu Gln Arg His Leu Glu Ile			
465	470	475	480
Leu Gln Gln Gln Leu Leu Gln Glu Gln Ala Met Leu Leu His Asp His			
485		490	495
Arg Arg Pro His Ala Gln Gln Pro Pro Pro Pro Gln Gln Gln Asp			
500	505	510	
Arg Ser Lys Pro Ser Phe His Ala Pro Glu Pro Lys Pro His Tyr Asp			
515	520	525	

Pro Ala Asp Arg Ala Arg Glu Val Gln Trp Ser His Leu Ala Ser Leu
530 535 540

Lys Asn Asn Val Ser Pro Val Ser Arg Ser His Ser Phe Ser Asp Pro
545 550 555 560

Ser Pro Lys Phe Ala His His Leu Arg Ser Gln Asp Pro Cys Pro
565 570 575

Pro Ser Arg Ser Glu Gly Leu Ser Gln Ser Ser Asp Ser Lys Ser Glu
580 585 590

Val Pro Glu Pro Thr Gln Lys Ala Trp Ser Arg Ser Asp Ser Asp Glu
595 600 605

Val Pro Pro Arg Val Pro Val Arg Thr Thr Ser Arg Ser Pro Val Leu
610 615 620

Ser Arg Arg Asp Ser Pro Leu Gln Gly Gly Gln Gln Asn Ser Gln
625 630 635 640

Ala Gly Gln Arg Asn Ser Thr Ser Ser Ile Glu Pro Arg Leu Leu Trp
645 650 655

Glu Arg Val Glu Lys Leu Val Pro Arg Pro Gly Ser Gly Ser Ser Ser
660 665 670

Gly Ser Ser Asn Ser Gly Ser Gln Pro Gly Ser His Pro Gly Ser Gln
675 680 685

Ser Gly Ser Gly Glu Arg Phe Arg Val Arg Ser Ser Ser Lys Ser Glu
690 695 700

Gly Ser Pro Ser Pro Arg Gln Glu Ser Ala Ala Lys Lys Pro Asp Asp
705 710 715 720

Lys Lys Glu Val Phe Arg Ser Leu Lys Pro Ala Gly Glu Val Asp Leu
725 730 735

Thr Ala Leu Ala Lys Glu Leu Arg Ala Val Glu Asp Val Arg Pro Pro
740 745 750

His Lys Val Thr Asp Tyr Ser Ser Ser Ser Glu Glu Ser Gly Thr Thr
755 760 765

Asp Glu Glu Glu Glu Asp Val Glu Gln Glu Gly Ala Asp Asp Ser Thr
770 775 780

Ser Gly Pro Glu Asp Thr Arg Ala Ala Ser Ser Pro Asn Leu Ser Asn
785 790 795 800

Gly Glu Thr Glu Ser Val Lys Thr Met Ile Val His Asp Asp Val Glu
805 810 815

89

Ser Glu Pro Ala Met Thr Pro Ser Lys Glu Gly Thr Leu Ile Val Arg
820 825 830

Gln Thr Gln Ser Ala Ser Ser Thr Leu Gln Lys His Lys Ser Ser Ser
835 840 845

Ser Phe Thr Pro Phe Ile Asp Pro Arg Leu Leu Gln Ile Ser Pro Ser
850 855 860

Ser Gly Thr Thr Val Thr Ser Val Val Gly Phe Ser Cys Asp Gly Leu
865 870 875 880

Arg Pro Glu Ala Ile Arg Gln Asp Pro Thr Arg Lys Gly Ser Val Val
885 890 895

Asn Val Asn Pro Thr Asn Thr Arg Pro Gln Ser Asp Thr Pro Glu Ile
900 905 910

Arg Lys Tyr Lys Lys Arg Phe Asn Ser Glu Ile Leu Cys Ala Ala Leu
915 920 925

Trp Gly Val Asn Leu Leu Val Gly Thr Glu Ser Gly Leu Met Leu Leu
930 935 940

Asp Arg Ser Gly Gln Gly Lys Val Tyr Pro Leu Ile Ser Arg Arg Arg
945 950 955 960

Phe Gln Gln Met Asp Val Leu Glu Gly Leu Asn Val Leu Val Thr Ile
965 970 975

Ser Gly Lys Lys Asp Lys Leu Arg Val Tyr Tyr Leu Ser Trp Leu Arg
980 985 990

Asn Lys Ile Leu His Asn Asp Pro Glu Val Glu Lys Lys Gln Gly Trp
995 1000 1005

Thr Thr Val Gly Asp Leu Glu Gly Cys Val His Tyr Lys Val Val Lys
1010 1015 1020

Tyr Glu Arg Ile Lys Phe Leu Val Ile Ala Leu Lys Ser Ser Val Glu
1025 1030 1035 1040

Val Tyr Ala Trp Ala Pro Lys Pro Tyr His Lys Phe Met Ala Phe Lys
1045 1050 1055

Ser Phe Gly Glu Leu Leu His Lys Pro Leu Leu Val Asp Leu Thr Val
1060 1065 1070

Glu Glu Gly Gln Arg Leu Lys Val Ile Tyr Gly Ser Cys Ala Gly Phe
1075 1080 1085

His Ala Val Asp Val Asp Ser Gly Ser Val Tyr Asp Ile Tyr Leu Pro
1090 1095 1100

90

Thr His Ile Gln Cys Ser Ile Lys Pro His Ala Ile Ile Ile Leu Pro
 1105 1110 1115 1120

Asn Thr Asp Gly Met Glu Leu Leu Val Cys Tyr Glu Asp Glu Gly Val
 1125 1130 1135

Tyr Val Asn Thr Tyr Gly Arg Ile Thr Lys Asp Val Val Leu Gln Trp
 1140 1145 1150

Gly Glu Met Pro Thr Ser Val Ala Tyr Ile Arg Ser Asn Gln Thr Met
 1155 1160 1165

Gly Trp Gly Glu Lys Ala Ile Glu Ile Arg Ser Val Glu Thr Gly His
 1170 1175 1180

Leu Asp Gly Val Phe Met His Lys Arg Ala Gln Arg Leu Lys Phe Leu
 1185 1190 1195 1200

Cys Gly Arg Asn Asp Lys Val Phe Phe Ser Ser Val Arg Ser Gly Gly
 1205 1210 1215

Ser Ser Gln Val Tyr Phe Met Thr Leu Gly Arg Thr Ser Leu Leu Ser
 1220 1225 1230

Trp

<210> 90
 <211> 982
 <212> PRT
 <213> SULU_ce

<400> 90

Met Ala Pro Ala Val Leu Asp Lys Pro Gly Val Ile Lys Asp Pro Ser
 1 5 10 15

Ile Ala Ala Leu Phe Ser Asn Lys Asp Pro Glu Gln Arg Tyr Gln Asp
 20 25 30

Leu Arg Glu Ile Gly His Gly Ser Phe Gly Ala Val Tyr Phe Ala Tyr
 35 40 45

Asp Lys Lys Asn Glu Gln Thr Val Ala Ile Lys Lys Met Asn Phe Ser
 50 55 60

Gly Lys Gln Ala Val Glu Lys Trp Asn Asp Ile Leu Lys Glu Val Ser
 65 70 75 80

Phe Leu Asn Thr Val Val His Pro His Ile Val Asp Tyr Lys Ala Cys
 85 90 95

Phe Leu Lys Asp Thr Thr Cys Trp Leu Val Met Glu Tyr Cys Ile Gly
 100 105 110

Ser Ala Ala Asp Ile Val Asp Val Leu Arg Lys Gly Met Arg Glu Val
115 120 125

Glu Ile Ala Ala Ile Cys Ser Gln Thr Leu Asp Ala Leu Arg Tyr Leu
130 135 140

His Ser Leu Lys Arg Ile His Arg Asp Ile Lys Ala Gly Asn Ile Leu
145 150 155 160

Leu Ser Asp His Ala Ile Val Lys Leu Ala Asp Phe Gly Ser Ala Ser
165 170 175

Leu Val Asp Pro Ala Gln Thr Phe Ile Gly Thr Pro Phe Phe Met Ala
180 185 190

Pro Glu Val Ile Leu Ala Met Asp Glu Gly His Tyr Thr Asp Arg Ala
195 200 205

Asp Ile Trp Ser Leu Gly Ile Thr Cys Ile Glu Leu Ala Glu Arg Arg
210 215 220

Pro Pro Leu Phe Ser Met Asn Ala Met Ser Ala Leu Tyr His Ile Ala
225 230 235 240

Gln Asn Asp Pro Pro Thr Leu Ser Pro Ile Asp Thr Ser Glu Gln Pro
245 250 255

Glu Trp Ser Leu Glu Phe Val Gln Phe Ile Asp Lys Cys Leu Arg Lys
260 265 270

Pro Ala Glu Glu Arg Met Ser Ala Glu Glu Cys Phe Arg His Pro Phe
275 280 285

Ile Gln Arg Ser Arg Pro Ser Asp Thr Ile Gln Glu Leu Ile Gln Arg
290 295 300

Thr Lys Asn Met Val Leu Glu Leu Asp Asn Phe Gln Tyr Lys Lys Met
305 310 315 320

Arg Lys Leu Met Tyr Leu Asp Glu Thr Glu Gly Lys Glu Gly Ser Glu
325 330 335

Gly Asn Gly Ala Ser Asp Asp Leu Asp Phe His Gly Asn Glu Ala Asn
340 345 350

Ser Ile Gly Arg Ala Gly Asp Ser Ala Ser Ser Arg Ser Ala Ser Leu
355 360 365

Thr Ser Phe Arg Ser Met Gln Ser Ser Gly Gly Ala Gly Leu Leu Val
370 375 380

Ser Thr Asn Thr Thr Gly Ala Met Asp Asn Val His Gly Ser Ser Gly
385 390 395 400

Tyr Gly Asn Gly Ser Ser Ser Thr Thr Ser Ser Ala Arg Arg Arg Pro
405 410 415

Pro Ile Pro Ser Gln Met Leu Ser Ser Thr Ser Thr Ser Gly Val Gly
420 425 430

Thr Met Pro Ser His Gly Ser Val Gly Ala Ser Ile Thr Ala Ile Ala
435 440 445

Val Asn Pro Thr Pro Ser Pro Ser Glu Pro Ile Pro Thr Ser Gln Pro
450 455 460

Thr Ser Lys Ser Glu Ser Ser Ile Leu Glu Thr Ala His Asp Asp
465 470 475 480

Pro Leu Asp Thr Ser Ile Arg Ala Pro Val Lys Asp Leu His Met Pro
485 490 495

His Arg Ala Val Lys Glu Arg Ile Ala Thr Leu Gln Asn His Lys Phe
500 505 510

Ala Thr Leu Arg Ser Gln Arg Ile Ile Asn Gln Glu Gln Glu Glu Tyr
515 520 525

Thr Lys Glu Asn Asn Met Tyr Glu Gln Met Ser Lys Tyr Lys His Leu
530 535 540

Arg Gln Ala His His Lys Glu Leu Gln Gln Phe Glu Glu Arg Cys Ala
545 550 555 560

Leu Asp Arg Glu Gln Leu Arg Val Lys Met Asp Arg Glu Leu Glu Gln
565 570 575

Leu Thr Thr Tyr Ser Lys Glu Lys Met Arg Val Arg Cys Ser Gln
580 585 590

Asn Asn Glu Leu Asp Lys Arg Lys Lys Asp Ile Glu Asp Gly Glu Lys
595 600 605

Lys Met Lys Lys Thr Lys Asn Ser Gln Asn Gln Gln Met Lys Leu
610 615 620

Tyr Ser Ala Gln Gln Leu Lys Glu Tyr Lys Tyr Asn Lys Glu Ala Gln
625 630 635 640

Lys Thr Arg Leu Arg Ser Leu Asn Met Pro Arg Ser Thr Tyr Glu Asn
645 650 655

Ala Met Lys Glu Val Lys Ala Asp Leu Asn Arg Val Lys Asp Ala Arg
660 665 670

Glu Asn Asp Phe Asp Glu Lys Leu Arg Ala Glu Leu Glu Asp Glu Ile
675 680 685

Val Arg Tyr Arg Arg Gln Gln Leu Ser Asn Leu His Gln Leu Glu Glu
690 695 700

Gln Leu Asp Asp Glu Asp Val Asn Val Gln Glu Arg Gln Met Asp Thr
705 710 715 720

Arg His Gly Leu Leu Ser Lys Gln His Glu Met Thr Arg Asp Leu Glu
725 730 735

Ile Gln His Leu Asn Glu Leu His Ala Met Lys Lys Arg His Leu Glu
740 745 750

Thr Gln His Glu Ala Glu Ser Ala Ser Gln Asn Glu Tyr Thr Gln Arg
755 760 765

Gln Gln Asp Glu Leu Arg Lys Lys His Ala Met Gln Ser Arg Gln Gln
770 775 780

Pro Arg Asp Leu Lys Ile Gln Glu Ala Gln Ile Arg Lys Gln Tyr Arg
785 790 795 800

Gln Val Val Lys Thr Gln Thr Arg Gln Phe Lys Leu Tyr Leu Thr Gln
805 810 815

Met Val Gln Val Val Pro Lys Asp Glu Gln Lys Glu Leu Thr Ser Arg
820 825 830

Leu Lys Gln Asp Gln Met Gln Lys Val Ala Leu Leu Ala Ser Gln Tyr
835 840 845

Glu Ser Gln Ile Lys Lys Met Val Gln Asp Lys Thr Val Lys Leu Glu
850 855 860

Ser Trp Gln Glu Asp Glu Gln Arg Val Leu Ser Glu Lys Leu Glu Lys
865 870 875 880

Glu Leu Glu Glu Leu Ile Ala Tyr Gln Lys Lys Thr Arg Ala Thr Leu
885 890 895

Glu Glu Gln Ile Lys Lys Glu Arg Thr Ala Leu Glu Glu Arg Ile Gly
900 905 910

Thr Arg Arg Ala Met Leu Glu Gln Lys Ile Ile Glu Glu Arg Glu Gln
915 920 925

Met Gly Glu Met Arg Arg Leu Lys Lys Glu Gln Ile Arg Asp Arg His
930 935 940

Ser Gln Glu Arg His Arg Leu Glu Asn His Phe Val Arg Thr Gly Ser
945 950 955 960

Thr Ser Arg Ser Ser Gly Gly Ile Ala Pro Gly Val Gly Asn Ser Ser
965 970 975

Ser Ile Gln Met Ala Met
980

<210> 91
<211> 842
<212> PRT
<213> KHS1_h

<400> 91

Leu Arg Pro Ala Ala Asp Ile Leu Arg Arg Asn Pro Gln Gln Asp Tyr
1 5 10 15

Glu Leu Val Gln Arg Val Gly Ser Gly Thr Tyr Gly Asp Val Tyr Lys
20 25 30

Ala Arg Asn Val His Thr Gly Glu Leu Ala Ala Val Lys Ile Ile Lys
35 40 45

Leu Glu Pro Gly Asp Asp Phe Ser Leu Ile Gln Gln Glu Ile Phe Met
50 55 60

Val Lys Glu Cys Lys His Cys Asn Ile Val Ala Tyr Phe Gly Ser Tyr
65 70 75 80

Leu Ser Arg Glu Lys Leu Trp Ile Cys Met Glu Tyr Cys Gly Gly
85 90 95

Ser Leu Gln Asp Ile Tyr His Val Thr Gly Pro Leu Ser Glu Leu Gln
100 105 110

Ile Ala Tyr Val Cys Arg Glu Thr Leu Gln Gly Leu Ala Tyr Leu His
115 120 125

Thr Lys Gly Lys Met His Arg Asp Ile Lys Gly Ala Asn Ile Leu Leu
130 135 140

Thr Asp His Gly Asp Val Lys Leu Ala Asp Phe Gly Val Ala Ala Lys
145 150 155 160

Ile Thr Ala Thr Ile Ala Lys Arg Lys Ser Phe Ile Gly Thr Pro Tyr
165 170 175

Trp Met Ala Pro Glu Val Ala Ala Val Glu Lys Asn Gly Gly Tyr Asn
180 185 190

Gln Leu Cys Asp Ile Trp Ala Val Gly Ile Thr Ala Ile Glu Leu Gly
195 200 205

Glu Leu Gln Pro Pro Met Phe Asp Leu His Pro Met Arg Ala Leu Phe
210 215 220

Leu Met Ser Lys Ser Asn Phe Gln Pro Pro Lys Leu Lys Asp Lys Thr
225 230 235 240

Lys Trp Ser Ser Thr Phe His Asn Phe Val Lys Ile Ala Leu Thr Lys
245 250 255

Asn Pro Lys Lys Arg Pro Thr Ala Glu Arg Leu Leu Thr His Thr Phe
 260 265 270

Val Ala Gln Pro Gly Leu Ser Arg Ala Leu Ala Val Glu Leu Leu Asp
 275 280 285

Lys Val Asn Asn Pro Asp Asn His Ala His Tyr Thr Glu Ala Asp Asp
 290 295 300

Asp Asp Phe Glu Pro His Ala Ile Ile Arg His Thr Ile Arg Ser Thr
 305 310 315 320

Asn Arg Asn Ala Arg Ala Glu Arg Thr Ala Ser Glu Ile Asn Phe Asp
 325 330 335

Lys Leu Gln Phe Glu Pro Pro Leu Arg Lys Glu Thr Glu Ala Arg Asp
 340 345 350

Glu Met Gly Leu Ser Ser Asp Pro Asn Phe Met Leu Gln Trp Asn Pro
 355 360 365

Phe Val Asp Gly Ala Asn Thr Gly Lys Ser Thr Ser Lys Arg Ala Ile
 370 375 380

Pro Pro Pro Leu Pro Pro Lys Pro Arg Ile Ser Ser Tyr Pro Glu Asp
 385 390 395 400

Asn Phe Pro Asp Glu Glu Lys Ala Ser Thr Ile Lys His Cys Pro Asp
 405 410 415

Ser Glu Ser Arg Ala Pro Gln Ile Leu Arg Arg Gln Ser Ser Pro Ser
 420 425 430

Cys Gly Pro Val Ala Glu Thr Ser Ser Ile Gly Asn Gly Asp Gly Ile
 435 440 445

Ser Lys Leu Met Ser Glu Asn Thr Glu Gly Ser Ala Gln Ala Pro Gln
 450 455 460

Leu Pro Arg Lys Asn Asp Lys Arg Asp Phe Pro Lys Pro Ala Ile Asn
 465 470 475 480

Gly Leu Pro Pro Thr Pro Lys Val Leu Met Gly Ala Cys Phe Ser Lys
 485 490 495

Val Phe Asp Gly Cys Pro Leu Lys Ile Asn Cys Ala Thr Ser Trp Ile
 500 505 510

His Pro Asp Thr Lys Asp Gln Tyr Ile Ile Phe Gly Thr Glu Asp Gly
 515 520 525

Ile Tyr Thr Leu Asn Leu Asn Glu Leu His Glu Ala Thr Met Glu Gln
 530 535 540

Leu Phe Pro Arg Lys Cys Thr Trp Leu Tyr Val Ile Asn Asn Thr Leu
545 550 555 560

Met Ser Leu Ser Glu Gly Lys Thr Phe Gln Leu Tyr Ser His Asn Leu
565 570 575

Ile Ala Leu Phe Glu His Ala Lys Lys Pro Gly Leu Ala Ala His Ile
580 585 590

Gln Thr His Arg Phe Pro Asp Arg Ile Leu Pro Arg Lys Phe Ala Leu
595 600 605

Thr Thr Lys Ile Pro Asp Thr Lys Gly Cys His Lys Cys Cys Ile Val
610 615 620

Arg Asn Pro Tyr Thr Gly His Lys Tyr Leu Cys Gly Ala Leu Gln Ser
625 630 635 640

Gly Ile Val Leu Leu Gln Trp Tyr Glu Pro Met Gln Lys Phe Met Leu
645 650 655

Ile Lys His Phe Asp Phe Pro Leu Pro Ser Pro Leu Asn Val Phe Glu
660 665 670

Met Leu Val Ile Pro Glu Gln Glu Tyr Pro Met Val Cys Val Ala Ile
675 680 685

Ser Lys Gly Thr Glu Ser Asn Gln Val Val Gln Phe Glu Thr Ile Asn
690 695 700

Leu Asn Ser Ala Ser Ser Trp Phe Thr Glu Ile Gly Ala Gly Ser Gln
705 710 715 720

Gln Leu Asp Ser Ile His Val Thr Gln Leu Glu Arg Asp Thr Val Leu
725 730 735

Val Cys Leu Asp Lys Phe Val Lys Ile Val Asn Leu Gln Gly Lys Leu
740 745 750

Lys Ser Ser Lys Lys Leu Ala Ser Glu Leu Ser Phe Asp Phe Arg Ile
755 760 765

Glu Ser Val Val Cys Leu Gln Asp Ser Val Leu Ala Phe Trp Lys His
770 775 780

Gly Met Gln Gly Lys Ser Phe Lys Ser Asp Glu Val Thr Gln Glu Ile
785 790 795 800

Ser Asp Glu Thr Arg Val Phe Arg Leu Leu Gly Ser Asp Arg Val Val
805 810 815

Val Leu Glu Ser Arg Pro Thr Glu Asn Pro Thr Ala His Ser Asn Leu
820 825 830

Tyr Ile Leu Ala Gly His Glu Asn Ser Tyr

835

840

<210> 92
<211> 911
<212> PRT
<213> LOK_m

<400> 92

Met Ala Phe Ala Asn Phe Arg Arg Ile Leu Arg Leu Ser Thr Phe Glu
1 5 10 15

Lys Arg Lys Ser Arg Glu Tyr Glu His Val Arg Arg Asp Leu Asp Pro
20 25 30

Asn Asp Val Trp Glu Ile Val Gly Glu Leu Gly Asp Gly Ala Phe Gly
35 40 45

Lys Val Tyr Lys Ala Lys Asn Lys Glu Thr Gly Ala Leu Ala Ala Ala
50 55 60

Lys Val Ile Glu Thr Lys Ser Glu Glu Leu Glu Asp Tyr Ile Val
65 70 75 80

Glu Ile Glu Ile Leu Ala Thr Cys Asp His Pro Tyr Ile Val Lys Leu
85 90 95

Leu Gly Ala Tyr Tyr Asp Gly Lys Leu Trp Ile Met Ile Glu Phe
100 105 110

Cys Pro Gly Gly Ala Val Asp Ala Ile Met Leu Glu Leu Asp Arg Gly
115 120 125

Leu Thr Glu Pro Gln Ile Gln Val Val Cys Arg Gln Met Leu Glu Ala
130 135 140

Leu Asn Phe Leu His Gly Lys Arg Ile Ile His Arg Asp Leu Lys Ala
145 150 155 160

Gly Asn Val Leu Met Thr Leu Glu Gly Asp Ile Arg Leu Ala Asp Phe
165 170 175

Gly Val Ser Ala Lys Asn Leu Lys Thr Leu Gln Lys Arg Asp Ser Phe
180 185 190

Ile Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Val Leu Cys Glu Thr
195 200 205

Met Lys Asp Ala Pro Tyr Asp Tyr Lys Ala Asp Ile Trp Ser Leu Gly
210 215 220

Ile Thr Leu Ile Glu Met Ala Gln Ile Glu Pro Pro His His Glu Leu
225 230 235 240

Asn Pro Met Arg Val Leu Leu Lys Ile Ala Lys Ser Asp Pro Pro Thr
245 250 255

Leu Leu Thr Pro Ser Lys Trp Ser Val Glu Phe Arg Asp Phe Leu Lys
260 265 270

Ile Ala Leu Asp Lys Asn Pro Glu Thr Arg Pro Ser Ala Ala Gln Leu
275 280 285

Leu Gln His Pro Phe Val Ser Arg Val Thr Ser Asn Lys Ala Leu Arg
290 295 300

Glu Leu Val Ala Glu Ala Lys Ala Glu Val Met Glu Glu Ile Glu Asp
305 310 315 320

Gly Arg Glu Asp Gly Glu Glu Asp Ala Val Asp Ala Val Pro Pro
325 330 335

Leu Val Asn His Thr Gln Asp Ser Ala Asn Val Thr Gln Pro Ser Leu
340 345 350

Asp Ser Asn Lys Leu Leu Gln Asp Ser Ser Thr Pro Leu Pro Pro Ser
355 360 365

Gln Pro Gln Glu Pro Val Asn Gly Pro Cys Ser Gln Pro Ser Gly Asp
370 375 380

Gly Pro Leu Gln Thr Thr Ser Pro Ala Asp Gly Leu Ser Lys Asn Asp
385 390 395 400

Asn Asp Leu Lys Val Pro Val Pro Leu Arg Lys Ser Arg Pro Leu Ser
405 410 415

Met Asp Ala Arg Ile Gln Met Asp Glu Glu Lys Gln Ile Pro Asp Gln
420 425 430

Asp Glu Asn Pro Ser Pro Ala Ala Ser Lys Ser Gln Lys Ala Asn Gln
435 440 445

Ser Arg Pro Asn Ser Ser Ala Leu Glu Thr Leu Gly Gly Glu Ala Leu
450 455 460

Thr Asn Gly Gly Leu Glu Leu Pro Ser Ser Val Thr Pro Ser His Ser
465 470 475 480

Lys Arg Ala Ser Asp Cys Ser Asn Leu Ser Thr Ser Glu Ser Met Asp
485 490 495

Tyr Gly Thr Ser Leu Ser Ala Asp Leu Ser Leu Asn Lys Glu Thr Gly
500 505 510

Ser Leu Ser Leu Lys Gly Ser Lys Leu His Asn Lys Thr Leu Lys Arg
515 520 525

Thr Arg Arg Phe Val Val Asp Gly Val Glu Val Ser Ile Thr Thr Ser
530 535 540

Lys Ile Ile Ser Glu Asp Glu Lys Lys Asp Glu Glu Met Arg Phe Leu
545 550 555 560

Arg Arg Gln Glu Leu Arg Glu Leu Arg Leu Leu Gln Lys Glu Glu His
565 570 575

Arg Asn Gln Thr Gln Leu Ser Ser Lys His Glu Leu Gln Leu Glu Gln
580 585 590

Met His Lys Arg Phe Glu Gln Glu Ile Asn Ala Lys Lys Lys Phe Tyr
595 600 605

Asp Val Glu Leu Glu Asn Leu Glu Arg Gln Gln Lys Gln Gln Val Glu
610 615 620

Lys Met Glu Gln Asp His Ser Val Arg Arg Lys Glu Glu Ala Lys Arg
625 630 635 640

Ile Arg Leu Glu Gln Asp Arg Asp Tyr Ala Lys Phe Gln Glu Gln Leu
645 650 655

Lys Gln Met Lys Lys Glu Val Lys Ser Glu Val Glu Lys Leu Pro Arg
660 665 670

Gln Gln Arg Lys Glu Ser Met Lys Gln Lys Met Glu Glu His Ser Gln
675 680 685

Lys Lys Gln Arg Leu Asp Arg Asp Phe Val Ala Lys Gln Lys Glu Asp
690 695 700

Leu Glu Leu Ala Met Arg Lys Leu Thr Thr Glu Asn Arg Arg Glu Ile
705 710 715 720

Cys Asp Lys Glu Arg Asp Cys Leu Ser Lys Lys Gln Glu Leu Leu Arg
725 730 735

Asp Arg Glu Ala Ala Leu Trp Glu Met Glu Glu His Gln Leu Gln Glu
740 745 750

Arg His Gln Leu Val Lys Gln Gln Leu Lys Asp Gln Tyr Phe Leu Gln
755 760 765

Arg His Asp Leu Leu Arg Lys His Glu Lys Glu Arg Glu Gln Met Gln
770 775 780

Arg Tyr Asn Gln Arg Met Met Glu Gln Leu Lys Val Arg Gln Gln Gln
785 790 795 800

Glu Lys Ala Arg Leu Pro Lys Ile Gln Arg Ser Asp Gly Glu Thr Arg
805 810 815

Met Ala Met Tyr Lys Lys Ser Leu His Ile Asn Gly Ala Gly Ser Ala

100

820 825 830

Ser Glu Gln Arg Glu Lys Ile Lys Gln Phe Ser Gln Gln Glu Glu Lys
835 840 845

Arg Gln Lys Ala Glu Arg Leu Gln Gln Gln Lys His Glu His Gln
850 855 860

Met Arg Asp Met Val Ala Gln Cys Glu Ser Asn Met Ser Glu Leu Gln
865 870 875 880

Gln Leu Gln Asn Glu Lys Cys Tyr Leu Leu Val Glu His Glu Thr Gln
885 890 895

Lys Leu Lys Ala Leu Asp Glu Ser His Asn Gln Ser Leu Lys Glu
900 905 910

<210> 93
<211> 545
<212> PRT
<213> PAK1_h

<400> 93

Met Ser Asn Asn Gly Leu Asp Ile Gln Asp Lys Pro Pro Ala Pro Pro
1 5 10 15

Met Arg Asn Thr Ser Thr Met Ile Gly Ala Gly Ser Lys Asp Ala Gly
20 25 30

Thr Leu Asn His Gly Ser Lys Pro Leu Pro Pro Asn Pro Glu Glu Lys
35 40 45

Lys Lys Lys Asp Arg Phe Tyr Arg Ser Ile Leu Pro Gly Asp Lys Thr
50 55 60

Asn Lys Lys Lys Glu Lys Glu Arg Pro Glu Ile Ser Leu Pro Ser Asp
65 70 75 80

Phe Glu His Thr Ile His Val Gly Phe Asp Ala Val Thr Gly Glu Phe
85 90 95

Thr Gly Met Pro Glu Gln Trp Ala Arg Leu Leu Gln Thr Ser Asn Ile
100 105 110

Thr Lys Ser Glu Gln Lys Lys Asn Pro Gln Ala Val Leu Asp Val Leu
115 120 125

Glu Phe Tyr Asn Ser Lys Lys Thr Ser Asn Ser Gln Lys Tyr Met Ser
130 135 140

101

Phe Thr Asp Lys Ser Ala Glu Asp Tyr Asn Ser Ser Asn Ala Leu Asn
145 150 155 160

Val Lys Ala Val Ser Glu Thr Pro Ala Val Pro Pro Val Ser Glu Asp
165 170 175

Glu Asp Asp Asp Asp Asp Ala Thr Pro Pro Pro Val Ile Ala Pro
180 185 190

Arg Pro Glu His Thr Lys Ser Val Tyr Thr Arg Ser Val Ile Glu Pro
195 200 205

Leu Pro Val Thr Pro Thr Arg Asp Val Ala Thr Ser Pro Ile Ser Pro
210 215 220

Thr Glu Asn Asn Thr Thr Pro Pro Asp Ala Leu Thr Leu Asn Thr Glu
225 230 235 240

Lys Gln Lys Lys Pro Lys Met Ser Asp Glu Glu Ile Leu Glu Lys
245 250 255

Leu Arg Ser Ile Val Ser Val Gly Asp Pro Lys Lys Tyr Thr Arg
260 265 270

Phe Glu Lys Ile Gly Gln Gly Ala Ser Gly Thr Val Tyr Thr Ala Met
275 280 285

Asp Val Ala Thr Gly Gln Glu Val Ala Ile Lys Gln Met Asn Leu Gln
290 295 300

Gln Gln Pro Lys Lys Glu Leu Ile Ile Asn Glu Ile Leu Val Met Arg
305 310 315 320

Glu Asn Lys Asn Pro Asn Ile Val Asn Tyr Leu Asp Ser Tyr Leu Val
325 330 335

Gly Asp Glu Leu Trp Val Val Met Glu Tyr Leu Ala Gly Gly Ser Leu
340 345 350

Thr Asp Val Val Thr Glu Thr Cys Met Asp Glu Gly Gln Ile Ala Ala
355 360 365

Val Cys Arg Glu Cys Leu Gln Ala Leu Glu Ser Leu His Ser Asn Gln
370 375 380

Val Ile His Arg Asp Ile Lys Ser Asp Asn Ile Leu Leu Gly Met Asp
385 390 395 400

Gly Ser Val Lys Leu Thr Asp Phe Gly Phe Cys Ala Gln Ile Thr Pro
405 410 415

Glu Gln Ser Lys Arg Ser Thr Met Val Gly Thr Pro Tyr Trp Met Ala
420 425 430

Pro Glu Val Val Thr Arg Lys Ala Tyr Gly Pro Lys Val Asp Ile Trp
435 440 445

Ser Leu Gly Ile Met Ala Ile Glu Met Ile Glu Gly Glu Pro Pro Tyr
450 455 460

Leu Asn Glu Asn Pro Leu Arg Ala Leu Tyr Leu Ile Ala Thr Asn Gly
465 470 475 480

Thr Pro Glu Leu Gln Asn Pro Glu Lys Leu Ser Ala Ile Phe Arg Asp
485 490 495

Phe Leu Asn Arg Cys Leu Glu Met Asp Val Glu Lys Arg Gly Ser Ala
500 505 510

Lys Glu Leu Leu Gln His Gln Phe Leu Lys Ile Ala Lys Pro Leu Ser
515 520 525

Ser Leu Thr Pro Leu Ile Ala Ala Ala Lys Glu Ala Thr Lys Asn Asn
530 535 540

His
545

<210> 94
<211> 506
<212> PRT
<213> PAK65_h

<400> 94

Met Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Ser Ala Asn
1 5 10 15

His Ser Leu Lys Pro Leu Pro Ser Val Pro Glu Glu Lys Lys Pro Arg
20 25 30

His Lys Ile Ile Ser Ile Phe Ser Gly Thr Glu Lys Gly Ser Lys Lys
35 40 45

Lys Glu Lys Glu Arg Pro Glu Ile Ser Pro Pro Ser Asp Phe Glu His
50 55 60

Thr Ile His Val Gly Phe Asp Thr Val Thr Gly Glu Phe Thr Gly Met
65 70 75 80

Pro Glu Gln Trp Ala Arg Leu Leu Gln Thr Ser Asn Ile Thr Lys Leu
85 90 95

Glu Gln Lys Lys Asn Pro Gln Ala Val Leu Asp Val Leu Lys Phe Tyr
100 105 110

Asp Ser Asn Thr Val Lys Gln Lys Tyr Leu Ser Phe Thr Pro Pro Glu
115 120 125

103

Lys Asp Gly Phe Pro Ser Gly Thr Pro Ala Leu Asn Ala Lys Gly Thr
130 135 140

Glu Ala Pro Ala Val Val Thr Glu Glu Glu Asp Asp Asp Glu Glu Thr
145 150 155 160

Ala Pro Pro Val Ile Ala Pro Arg Pro Asp His Thr Lys Ser Ile Tyr
165 170 175

Thr Arg Ser Val Ile Asp Pro Val Pro Ala Pro Val Gly Asp Ser His
180 185 190

Val Asp Gly Ala Ala Lys Ser Leu Asp Lys Gln Lys Lys Lys Thr Lys
195 200 205

Met Thr Asp Glu Glu Ile Met Glu Lys Leu Arg Thr Ile Val Ser Ile
210 215 220

Gly Asp Pro Lys Lys Lys Tyr Thr Arg Tyr Glu Lys Ile Gly Gln Gly
225 230 235 240

Ala Ser Gly Thr Val Phe Thr Ala Thr Asp Val Ala Leu Gly Gln Glu
245 250 255

Val Ala Ile Lys Gln Ile Asn Leu Gln Lys Gln Pro Lys Lys Glu Leu
260 265 270

Ile Ile Asn Glu Ile Leu Val Met Lys Glu Leu Lys Asn Pro Asn Ile
275 280 285

Val Asn Phe Leu Asp Ser Tyr Leu Val Gly Asp Glu Leu Phe Val Val
290 295 300

Met Glu Tyr Leu Ala Gly Arg Ser Leu Thr Asp Val Val Thr Glu Thr
305 310 315 320

Cys Met Asp Glu Ala Gln Ile Ala Ala Val Cys Arg Glu Cys Leu Gln
325 330 335

Ala Leu Glu Phe Leu His Ala Asn Gln Val Ile His Arg Asp Ile Lys
340 345 350

Ser Asp Asn Val Leu Leu Gly Met Glu Gly Ser Val Lys Leu Thr Asp
355 360 365

Phe Gly Phe Cys Ala Gln Ile Thr Pro Glu Gln Ser Lys Arg Ser Thr
370 375 380

Met Val Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Val Thr Arg Lys
385 390 395 400

Ala Tyr Gly Pro Lys Val Asp Ile Trp Ser Leu Gly Ile Met Ala Ile
405 410 415

Glu Met Val Glu Gly Glu Pro Pro Tyr Leu Asn Glu Asn Pro Leu Arg

104

420

425

430

Ala Leu Tyr Leu Ile Ala Thr Asn Gly Thr Pro Glu Leu Gln Asn Pro
435 440 445

Glu Lys Leu Ser Pro Ile Phe Arg Asp Phe Leu Asn Arg Cys Leu Glu
450 455 460

Met Asp Val Glu Lys Arg Gly Ser Ala Lys Glu Leu Leu Gln His Pro
465 470 475 480

Phe Leu Lys Leu Ala Lys Pro Leu Ser Ser Leu Thr Pro Leu Ile Met
485 490 495

Ala Ala Lys Glu Ala Met Lys Ser Asn Arg
500 505

<210> 95
<211> 544
<212> PRT
<213> PAK3_m

<400> 95

Met Ser Asp Ser Leu Asp Asn Glu Glu Lys Pro Pro Ala Pro Pro Leu
1 5 10 15

Arg Met Asn Ser Asn Asn Arg Asp Ser Ser Ala Leu Asn His Ser Ser
20 25 30

Lys Pro Leu Pro Met Ala Pro Glu Glu Lys Asn Lys Lys Ala Arg Leu
35 40 45

Arg Ser Ile Phe Pro Gly Gly Asp Lys Thr Asn Lys Lys Lys Glu
50 55 60

Lys Glu Arg Pro Glu Ile Ser Leu Pro Ser Asp Phe Glu His Thr Ile
65 70 75 80

His Val Gly Phe Asp Ala Val Thr Gly Glu Phe Thr Gly Ile Pro Glu
85 90 95

Gln Trp Ala Arg Leu Leu Gln Thr Ser Asn Ile Thr Lys Leu Glu Gln
100 105 110

Lys Lys Asn Pro Gln Ala Val Leu Asp Val Leu Lys Phe Tyr Asp Ser
115 120 125

Lys Glu Thr Val Asn Asn Gln Lys Tyr Met Ser Phe Thr Ser Gly Asp
130 135 140

Lys Ser Ala His Gly Tyr Ile Ala Ala His Gln Ser Asn Thr Lys Thr
145 150 155 160

105

Gly Ser Glu Pro Pro Leu Ala Pro Pro Val Ser Glu Glu Glu Asp Glu
165 170 175

Glu Glu Glu Glu Glu Asp Asp Asn Glu Pro Pro Pro Val Ile Ala
180 185 190

Pro Arg Pro Glu His Thr Lys Ser Ile Tyr Thr Arg Ser Val Val Glu
195 200 205

Ser Ile Ala Ser Pro Ala Ala Pro Asn Lys Glu Asp Ile Pro Pro Ser
210 215 220

Ala Glu Asn Ala Asn Ser Thr Thr Leu Tyr Arg Asn Thr Asp Arg Gln
225 230 235 240

Arg Lys Lys Ser Lys Met Thr Asp Glu Glu Ile Leu Glu Lys Leu Arg
245 250 255

Ser Ile Val Ser Val Gly Asp Pro Lys Lys Tyr Thr Arg Leu Glu
260 265 270

Lys Ile Gly Gln Gly Ala Ser Gly Thr Val Tyr Thr Ala Leu Asp Ile
275 280 285

Ala Thr Gly Gln Glu Val Ala Ile Lys Gln Met Asn Leu Gln Gln Gln
290 295 300

Pro Lys Lys Glu Leu Ile Ile Asn Glu Ile Leu Val Met Arg Glu Asn
305 310 315 320

Lys Asn Pro Asn Ile Val Asn Tyr Leu Asp Ser Tyr Leu Val Gly Asp
325 330 335

Glu Leu Trp Val Val Met Glu Tyr Leu Ala Gly Gly Ser Leu Thr Asp
340 345 350

Val Val Thr Glu Thr Cys Met Asp Val Gly Gln Ile Ala Ala Val Cys
355 360 365

Arg Glu Cys Leu Gln Ala Leu Asp Phe Leu His Ser Asn Gln Val Ile
370 375 380

His Arg Asp Ile Lys Ser Asp Asn Ile Leu Leu Gly Met Asp Gly Ser
385 390 395 400

Val Lys Leu Thr Asp Phe Gly Phe Cys Ala Gln Ile Thr Pro Glu Gln
405 410 415

Ser Lys Arg Ser Thr Met Val Gly Thr Pro Tyr Trp Met Ala Pro Glu
420 425 430

Val Val Thr Arg Lys Ala Tyr Gly Pro Lys Val Asp Ile Trp Ser Leu
435 440 445

Gly Ile Met Ala Ile Glu Met Val Glu Gly Glu Pro Pro Tyr Leu Asn

450	455	460
Glu Asn Pro Leu Arg Ala Leu Tyr Leu Ile Ala Thr Asn Gly Thr Pro		
465	470	475
Glu Leu Gln Asn Pro Glu Arg Leu Ser Ala Val Phe His Asp Phe Leu		
485	490	495
Asn Arg Cys Leu Glu Met Asp Val Asp Arg Arg Gly Ser Ala Lys Glu		
500	505	510
Leu Leu Gln His Pro Phe Leu Lys Leu Ala Lys Pro Leu Ser Ser Leu		
515	520	525
Thr Pro Leu Ile Ile Ala Ala Lys Glu Ala Ile Lys Asn Ser Ser Arg		
530	535	540

<400> 96
<210> 96
<211> 2110
<212> DNA
<213> Full Length Mammalian (Human) STLK5

<400> 96

ggccaagacg gtcggggctg cttgctaact ccaggaacag gtttaagttt ttgaaactga	60
atgttgtcta cacagtagga actcatgtca tttcttgtaa gtaaaccaga gcgaatcagg	120
cgggtggctc cgaaaagggtt cattgttag gccttaagag atttggaaact atttggagac	180
caatgatgcg agctcagagt caatagcatc cttctctaaa caggaggtca tgtagtagctt	240
tctgcccagag ggagggtgtt acgagctgtc cactgtgata ggcaaaggat ttgaggacct	300
gtgactgtg aatctagcaa ggtacaaaacc aacaggagag tacgtgactg tacggaggat	360
taaccttagaa gcttgttcca atgagatggt aacattcttg cagggcgagc tgcattgtctc	420
caaactcttc aaccatccca atatcggtcc atatcgagcc acttttattt cagacaatga	480
gctgtgggtt gtcacatcat tcattgcata cggttctgca aaagatctca tctgtacaca	540
cttcattgtat ggcattgtat agctggcgat tgcttacatc ctgcaggggg tgctgaaggc	600
cctcgactac atccaccaca tggatgtatgt acacaggagt gtcaagccca gccacatcct	660
gatctctgtg gatggaaagg tctacctgtc tggtttgcgc agcaacctca gcatgataag	720
ccatgggcag cggcagcgag tggtcacga tttcccaag tacagtgtca agttctgcc	780
gtggctcagc cccgagggtcc tccagcagaa tctccagggt tatgtatgcca agtctgacat	840
ctacagtgtg ggaatcacag cctgtgaact gccaacccgc catgtcccct ttaaggatat	900
gcctgccacc cagatgtgc tagagaaact gaacggcaca gtgcctgccc tggatatac	960
cagcaccatc cccgctgagg agctgaccat gagcccttcg cgctcagtgg ccaactctgg	1020
cctgagtgtac agcctgacca ccagcacccc cccgcctcc aacggtgact cggccctccca	1080
ccccctaccac cgaaccttct ccccccactt ccaccactt gtggagcagt gccttcagcg	1140
caacccggat gcccggccca gtgcacac cctcttgcac cacttttct tcaagcagat	1200
caagcgacgt gcctcagagg ctttgcgcga attgcttcgt cctgtcaccgc ccatcaccaa	1260
ttttgagggc agccagtctc aggaccacag tggatctt ggcctggtaa caaacctggaa	1320
agagctggag gtggacgatt gggaggctgt agcctctgca aactgtgcgc attctccagc	1380
caggatgtca gaggccaccc agaggccctt cctgaggccc ggcacattc cccgcctcc	1440
gggcagattg ggtagaaagg acattttcc agggaaagtgt actgtcact gattggaaaa	1500
aaaaatctcg gagagatact tcactgtcc aaggctttt agacacaagg gaatctcaac	1560
aaccaggat caggagggtc caaaggccgac attcccaagtc ctgtgagctc aggtgaccc	1620

ctccgcagaa gagagatgct gctctggccc tgggagctga atccaaggcc cagggtttgg	1680
ctccttaaac ccgaggaccg ccaccccttc ccagtgcctt cgaccaggct cattctattt	1740
aactttgctc tcagatgcct cagatgctat aggtcaagtga aaggcaagt agtaagctgc	1800
ctgcctccct tccctcagac ctctccctca taattccaga gaagggcatt tctgtctttt	1860
taagcacaga ctaaggctgg aacagtcctt ctttatccctt cttctggctt gggccctgac	1920
acctaagtct ttccccacggt ttatgtgtgt gcctcattcc tttcccaacca agaatccatc	1980
ttagcgctc ctgccagctg ccctggctt ttctccaaagg gccatcagtg tcttgcttag	2040
cttgagggtt taagtccctta tgctgtgtta gtttcgttgtt cagaacaaat taaaatttc	2100
agagacgctg	2110

<210> 97

<211> 373

<212> PRT

<213> Full Length Mammalian (Human) STLK 5

<400> 97

Met Ser Ser Phe Leu Pro Glu Gly Gly Cys Tyr Glu Leu Leu Thr Val			
1	5	10	15

Ile Gly Lys Gly Phe Glu Asp Leu Met Thr Val Asn Leu Ala Arg Tyr			
20	25	30	

Lys Pro Thr Gly Glu Tyr Val Thr Val Arg Arg Ile Asn Leu Glu Ala			
35	40	45	

Cys Ser Asn Glu Met Val Thr Phe Leu Gln Gly Glu Leu His Val Ser			
50	55	60	

Lys Leu Phe Asn His Pro Asn Ile Val Pro Tyr Arg Ala Thr Phe Ile			
65	70	75	80

Ala Asp Asn Glu Leu Trp Val Val Thr Ser Phe Met Ala Tyr Gly Ser			
85	90	95	

Ala Lys Asp Leu Ile Cys Thr His Phe Met Asp Gly Met Asn Glu Leu			
100	105	110	

Ala Ile Ala Tyr Ile Leu Gln Gly Val Leu Lys Ala Leu Asp Tyr Ile			
115	120	125	

His His Met Gly Tyr Val His Arg Ser Val Lys Ala Ser His Ile Leu			
130	135	140	

Ile Ser Val Asp Gly Lys Val Tyr Leu Ser Gly Leu Arg Ser Asn Leu			
145	150	155	160

Ser Met Ile Ser His Gly Gln Arg Gln Arg Val Val His Asp Phe Pro			
165	170	175	

Lys Tyr Ser Val Lys Val Leu Pro Trp Leu Ser Pro Glu Val Leu Gln			
180	185	190	

Gln Asn Leu Gln Gly Tyr Asp Ala Lys Ser Asp Ile Tyr Ser Val Gly
 195 200 205

Ile Thr Ala Cys Glu Leu Ala Asn Gly His Val Pro Phe Lys Asp Met
 210 215 220

Pro Ala Thr Gln Met Leu Leu Glu Lys Leu Asn Gly Thr Val Pro Cys
 225 230 235 240

Leu Leu Asp Thr Ser Thr Ile Pro Ala Glu Glu Leu Thr Met Ser Pro
 245 250 255

Ser Arg Ser Val Ala Asn Ser Gly Leu Ser Asp Ser Leu Thr Thr Ser
 260 265 270

Thr Pro Arg Pro Ser Asn Gly Asp Ser Pro Ser His Pro Tyr His Arg
 275 280 285

Thr Phe Ser Pro His Phe His His Phe Val Glu Gln Cys Leu Gln Arg
 290 295 300

Asn Pro Asp Ala Arg Pro Ser Ala Ser Thr Leu Leu Asn His Ser Phe
 305 310 315 320

Phe Lys Gln Ile Lys Arg Arg Ala Ser Glu Ala Leu Pro Glu Leu Leu
 325 330 335

Arg Pro Val Thr Pro Ile Thr Asn Phe Glu Gly Ser Gln Ser Gln Asp
 340 345 350

His Ser Gly Ile Phe Gly Leu Val Thr Asn Leu Glu Glu Leu Glu Val
 355 360 365

Asp Asp Trp Glu Phe
 370

<210> 98
 <211> 2001
 <212> DNA
 <213> Mammalian (Human) STLK6

<400> 98

aaggaagata aaacaaaagc cttcttggaa atagatggat ttttgtca ct ttctgtgtga
 actaaagtga ttcaatgtct cttttggatt gcttctgcac ttcaagaaca caagttgaat 120
 cactcagacc taaaaaacag tctgaaacca gtatccatca atacttggtt gatgagccaa
 ccctttctg gtcacgtcca tccacttagag ccagtgaagt actatgttcc accaacgttt
 ctcactatga gtcccaagta gaaataggaa gaggatttga caacttgact tctgtccatc 240
 ttgcacggca tactcccacg ggaacactgg taactataaa aattacaaaat ctggaaaact 300
 gcaatgaaga acgcctgaaa gcttacaga aagccgtgat tctatcccac ttttccggc 360
 420

atcccaatat tacaacttat tggacagttt tcactgttgg cagctggctt tgggttattt	480
ctccatttat ggcttatggt tcagcaagtc aactcttgag gacctattt cctgaaggaa	540
ttagtggaaac ttataataaga aacattctct ttggagccgt gagagggtt aactatctgc	600
accaaaatgg ctgtattcac aggagtatta aagccagcca tattcctcatt tctgggtatg	660
gccttagtgac cctctctggc ctttcccate tgcatacggtt ggttaagcat ggacagaggc	720
atagggctgt gtatgatttc ccacagttca gcacatcagt gcagccgtgg ctgagtccag	780
aactactgag acaggattta catgggtata atgtgaagtc agatatttac agtgttggaa	840
ttacagcatg tgaatttagcc agtgggcagg tgcctttcca ggacatgcat agaactcaga	900
tgctgttaca gaaactgaaa ggtccttcatt atagcccatt ggatatcagt atttccctc	960
aatcagaatc cagaatgaaa aattccctgt caggtgtaga ctctggatt ggagaaagtg	1020
tgcttgctc cagtggaaact cacacagtaa atagtgaccg attacacaca ccatcctcaa	1080
aaaccttctc tccgccttc tttagcttgg tacagctctg tttgcaacaa gatcctgaga	1140
aaaggccatc agcaaggcgt ttattgtccc atgttttctt caaacagatg aaagaagaaa	1200
gccaggattc aatactttca ctgttcctc ctgcttataa caagccatca atatcattgc	1260
ctccagtggtt accttggact gagccagaat gtgatttcc tgatgaaaaa gactcatact	1320
ggaaattcta gggctgccaa atcattttat gtcctatata ctggacactt tctccttgct	1380
gcttttctt ctgtatttctt aggtacaaat accagaatta tacttgaaaa tacagttgg	1440
gcactggaga atctattttt taaaaccact ctgttcaag gggcaccagt ttgttagtccc	1500
tctgtttcgc acagagtaact atgacaagga aacatcagaa ttactaatct agctagtgtc	1560
atttattctg gaattttttt ctaagctgtg actaactctt ttatctctc aatataattt	1620
ttgagccagt taattttttt cagtagtttgc ctgtccctt ggaatggcc ctcagaggac	1680
agtgccttcca agtacatctt ctcccagatt ctctggcctt ttatgagc tattgttaaa	1740
ccaacagggct agtttatctt acatcagacc ctttctggt agagggaaaa tggggcttgc	1800
ttccctttt cttctgttaa tacttatggt aacacctaac tgagcctcac tcacattaaa	1860
tgattcactt gaaatatata cagaaattgt aatttgctt ttttaaaaaa agggggctaa	1920
agtaacactt tcctacttat gtaaattata gatcctaaat tcacgcaccc cgtggagct	1980
caataaaagat ttactgaattt g	2001

<210> 99

<211> 373

<212> PRT

<213> Mammalian (Human) STLK6

<400> 99

Met Ser Ser Phe Leu Pro Glu Gly Gly Cys Tyr Glu Leu Leu Thr Val			
1	5	10	15

Ile Gly Lys Gly Phe Glu Asp Leu Met Thr Val Asn Leu Ala Arg Tyr		
20	25	30

Lys Pro Thr Gly Glu Tyr Val Thr Val Arg Arg Ile Asn Leu Glu Ala		
35	40	45

Cys Ser Asn Glu Met Val Thr Phe Leu Gln Gly Glu Leu His Val Ser		
50	55	60

Lys Leu Phe Asn His Pro Asn Ile Val Pro Tyr Arg Ala Thr Phe Ile			
65	70	75	80

Ala Asp Asn Glu Leu Trp Val Val Thr Ser Phe Met Ala Tyr Gly Ser		
85	90	95

110

Ala Lys Asp Leu Ile Cys Thr His Phe Met Asp Gly Met Asn Glu Leu
100 105 110

Ala Ile Ala Tyr Ile Leu Gln Gly Val Leu Lys Ala Leu Asp Tyr Ile
115 120 125

His His Met Gly Tyr Val His Arg Ser Val Lys Ala Ser His Ile Leu
130 135 140

Ile Ser Val Asp Gly Lys Val Tyr Leu Ser Gly Leu Arg Ser Asn Leu
145 150 155 160

Ser Met Ile Ser His Gly Gln Arg Gln Arg Val Val His Asp Phe Pro
165 170 175

Lys Tyr Ser Val Lys Val Leu Pro Trp Leu Ser Pro Glu Val Leu Gln
180 185 190

Gln Asn Leu Gln Gly Tyr Asp Ala Lys Ser Asp Ile Tyr Ser Val
195 200 205

Ile Thr Ala Cys Glu Leu Ala Asn Gly His Val Pro Phe Lys Asp Met
210 215 220

Pro Ala Thr Gln Met Leu Leu Glu Lys Leu Asn Gly Thr Val Pro
225 230 235 240

Leu Leu Asp Thr Ser Thr Ile Pro Ala Glu Glu Leu Thr Met Ser Pro
245 250 255

Ser Arg Ser Val Ala Asn Ser Gly Leu Ser Asp Ser Leu Thr Thr
260 265 270

Thr Pro Arg Pro Ser Asn Gly Asp Ser Pro Ser His Pro Tyr His Arg
275 280 285

Thr Phe Ser Pro His Phe His His Phe Val Glu Gln Cys Leu Gln
290 295 300

Asn Pro Asp Ala Arg Pro Ser Ala Ser Thr Leu Leu Asn His Ser Phe
305 310 315 320

Phe Lys Gln Ile Lys Arg Arg Ala Ser Glu Ala Leu Pro Glu Leu
325 330 335

Arg Pro Val Thr Pro Ile Thr Asn Phe Glu Gly Ser Gln Ser Gln Asp
340 345 350

His Ser Gly Ile Phe Gly Leu Val Thr Asn Leu Glu Glu Leu Glu
355 360 365

Asp Asp Trp Glu Phe

<210> 100
<211> 311
<212> DNA
<213> Mammalian (Human) STLK7

<400> 100

tcaacaggga cgattacgag ctgcaggagg tgatcgggag tggagcaact gctgttagtcc
aaggcagctta ttgtccccct aaaaaggaga aagtggcaat caaacggata aaccttgaga 180
aatgtcaaac tagcatggat gaactcctga aagaaaattca agccatgagt caatgccatc 240
atcctaataat tgttatcttac tacacatctt ttgtggtaaa agatgagctg tggcttgtca 300
tgaagctgct aagtggaggt tctgttctgg atattattaa gcacattgtg gcaaaagggg 311
aacacaaaaag t

<210> 101
<211> 103
<212> PRT
<213> Mammalian (Human) STLK7

<400> 101

Asn Arg Asp Asp Tyr Glu Leu Gln Glu Val Ile Gly Ser Gly Ala Thr
1 5 10 15

Ala Val Val Gln Ala Ala Tyr Cys Ala Pro Lys Lys Glu Lys Val Ala
20 25 30

Ile Lys Arg Ile Asn Leu Glu Lys Cys Gln Thr Ser Met Asp Glu Leu
35 40 45

Leu Lys Glu Ile Gln Ala Met Ser Gln Cys His His Pro Asn Ile Val
50 55 60

Ser Tyr Tyr Thr Ser Phe Val Val Lys Asp Glu Leu Trp Leu Val Met
65 70 75 80

Lys Leu Leu Ser Gly Gly Ser Val Leu Asp Ile Ile Lys His Ile Val
85 90 95

Ala Lys Gly Glu His Lys Ser
100

<210> 102
<211> 2806

<212> DNA
 <213> Full Length Mammalian (Human) PAK5

<400> 102

cgggagtgtc	cgcgggttgt	gcgggtcaag	agagctgaag	gaggcgcgag	ggcgccggagt	60
tccaggccga	gcagtttaggc	cgcgagcgac	tgcggcgccg	agccatgag	taaccggaaag	120
cccctagagg	agtggtcacc	tgcctgaggg	cacttctgtc	ccaccagcat	cagaccaggc	180
cgcaccgagt	ccccggcacc	atgtttggga	agaggaagaa	gcgggtggag	atctccgcgc	240
cgtccaactt	cgagcacccgc	gtgcacacgg	gcttcgacca	gcacagcag	aagttcacgg	300
ggctgccccc	ccagtggcag	agcctgatcg	aggagtgcggc	tcgcccggccc	aagccctcg	360
tcgaccccg	ctgcatcacc	tccatccagc	ccggggccccc	caagaccatc	gtgcggggca	420
gcaaagggtgc	caaagatggg	gccctcacgc	tgctgcttgg	cgagttttag	aacatgtcgg	480
tgacacgctc	caactccctg	cgagagagaca	gcccggccgc	gcccggccgt	gcccggccagg	540
aaaatgggat	gccagaggag	ccggccacca	cgcccgagagg	gggcccaggg	aaggcaggca	600
gccgaggccg	gttcggcggt	cacagcgagg	cggggtggcg	cagtggtgac	aggcgacggg	660
cgggggccaga	gaagaggccc	aagtcttcca	gggaggggttc	aggggttccc	caggagtct	720
ccccggacaa	acgccccctc	tccggccctg	atgtcgccac	cccccagct	gctggcttgg	780
ccagtggggc	gaaactggca	gttggccggc	ccttaaacac	ctaccggagg	gctgacacgg	840
accaccatc	ccgggggtgcc	cagggggagc	ctcatgacgt	ggcccttaac	ggccatcag	900
cggggggcct	ggccatcccc	cagtctctt	cctcttcctc	ccggeetccc	acccgagccc	960
gaggtggccc	cagcccttgg	gtgctggac	cccacgeetc	agagccccag	ctggcccttc	1020
cagcctgcac	ccccggccgc	cctgtgttc	ctggggccccc	tggcccccgc	tcaccacagc	1080
gggagccaca	gcgagtatcc	catgagcagt	tccgggctgc	cctgcagctg	gtggtgacc	1140
caggcgaccc	ccgctcttac	ctggacaact	tcatcaagat	tggcgagggc	tccacggca	1200
tcgtgtcat	cgccaccgtg	cgcagctcg	gcaagcttgt	ggccgtcaag	aagatggacc	1260
tcgcgaagca	gcagaggcgc	gagctgtct	tcaacgaggt	gtaatcatg	agggactacc	1320
agcacgagaa	tgtggtgag	atgtacaaca	gttaccttgt	gggggacgag	ctctgggtgg	1380
tcatggagtt	ccttggagga	ggccatctca	ccgacatctgt	caccacacc	aggatgaacg	1440
aggagcagat	cgcggccgt	tgccttgcag	tgctgcaggg	cctgtcggt	ctccacggccc	1500
agggcgtcat	ccaccgggac	atcaagagcg	actcgatct	gctgaccat	gatggcagg	1560
tgaagctgtc	agactttggg	ttctgcgtcc	aggtagccaa	ggaagtgc	cgaaggaagt	1620
cgctggtcgg	cacggccctac	tggatggccc	cagagctcat	ctccgcctt	ccctacgggc	1680
cagaggtaga	catctggtcg	ctgggataa	tggtattga	gatggtgac	ggagagcccc	1740
cctacttcaa	cgagccaccc	ctcaaaagcca	tgaagatgtat	tcggacaac	ctgccacccc	1800
gactgaagaa	cctgcacaag	gtgtcgccat	ccctgaaggg	cttcctggac	cgcctgtgg	1860
tgcgagaccc	tgcggccgg	gccacggcag	ccgagctgt	gaagcacc	tccctggcca	1920
aggcaggccc	gcgtggccage	atcgcccc	tcatgcgcca	gaaccgcacc	agatgaggcc	1980
cagcgcctt	ccccctcaacc	aaagagcccc	ccgggtcac	ccccccccca	ctgaggccag	2040
tagggggcca	ggccctccac	tcctcccagc	ccgggagatg	ctccgcgtgg	caccacccctc	2100
cttgcgtggg	gtagatgaga	ccctactact	gaactccagt	tttgcgttgc	tgacttttag	2160
aaaaacacag	ggactcgtgg	gagcaagcga	ggctcccagg	accccccaccc	tctggacag	2220
gccctcccc	atgttcttct	gtctccagga	aggcagcg	ccctccatc	actggaaatc	2280
tgcagtgggg	gtcgctgggg	gtggagagaa	actaagagg	tgaatgtt	tgagtgtgt	2340
cacgcgtgt	atgtgtcat	tgtgtgtgt	tgcaaagg	cagccacccc	gtccctccagc	2400
ccgcaagggg	tgtctggcgc	ttgcctgac	acccagcccc	ctctccccct	gagccattgt	2460
gggggtcgat	catgaatgtc	cgaagagtgg	cctttcccg	taaccgtcg	ccccctttct	2520
gtggctggat	ggggagacag	gtcaggccc	cccacccctc	ccagccccgt	cagcaaatga	2580
ctactgcacc	tggacagcc	cctctttct	agaagtctat	ttatattgtc	atttataac	2640
actctagccc	ctggcccttat	tggggacag	atggccctg	tcctgcgggg	tggccctggc	2700
agaaccactg	cctgaagaac	cagttcctg	cccggtcagc	gcagccccag	cccgccccacc	2760
cctgcctcga	gttagttta	caattaaaac	attgtctgt	tttgt		2806

<210> 103
<211> 591
<212> PRT
<213> Full Length Mammalian (Human) PAK5hu

<400> 103

Met Phe Gly Lys Arg Lys Lys Arg Val Glu Ile Ser Ala Pro Ser Asn
1 5 10 15

Phe Glu His Arg Val His Thr Gly Phe Asp Gln His Glu Gln Lys Phe
20 25 30

Thr Gly Leu Pro Arg Gln Trp Gln Ser Leu Ile Glu Glu Ser Ala Arg
35 40 45

Arg Pro Lys Pro Leu Val Asp Pro Ala Cys Ile Thr Ser Ile Gln Pro
50 55 60

Gly Ala Pro Lys Thr Ile Val Arg Gly Ser Lys Gly Ala Lys Asp Gly
65 70 75 80

Ala Leu Thr Leu Leu Leu Asp Glu Phe Glu Asn Met Ser Val Thr Arg
85 90 95

Ser Asn Ser Leu Arg Arg Asp Ser Pro Pro Pro Pro Ala Arg Ala Arg
100 105 110

Gln Glu Asn Gly Met Pro Glu Glu Pro Ala Thr Thr Ala Arg Gly Gly
115 120 125

Pro Gly Lys Ala Gly Ser Arg Gly Arg Phe Ala Gly His Ser Glu Ala
130 135 140

Gly Gly Gly Ser Gly Asp Arg Arg Arg Ala Gly Pro Glu Lys Arg Pro
145 150 155 160

Lys Ser Ser Arg Glu Gly Ser Gly Gly Pro Gln Glu Ser Ser Arg Asp
165 170 175

Lys Arg Pro Leu Ser Gly Pro Asp Val Gly Thr Pro Gln Pro Ala Gly
180 185 190

Leu Ala Ser Gly Ala Lys Leu Ala Ala Gly Arg Pro Phe Asn Thr Tyr
195 200 205

Pro Arg Ala Asp Thr Asp His Pro Ser Arg Gly Ala Gln Gly Glu Pro
210 215 220

His Asp Val Ala Pro Asn Gly Pro Ser Ala Gly Gly Leu Ala Ile Pro
225 230 235 240

Gln Ser Ser Ser Ser Ser Arg Pro Pro Thr Arg Ala Arg Gly Ala
245 250 255

Pro Ser Pro Gly Val Leu Gly Pro His Ala Ser Glu Pro Gln Leu Ala
260 265 270

Pro Pro Ala Cys Thr Pro Ala Ala Pro Ala Val Pro Gly Pro Pro Gly
275 280 285

Pro Arg Ser Pro Gln Arg Glu Pro Gln Arg Val Ser His Glu Gln Phe
290 295 300

Arg Ala Ala Leu Gln Leu Val Val Asp Pro Gly Asp Pro Arg Ser Tyr
305 310 315 320

Leu Asp Asn Phe Ile Lys Ile Gly Glu Gly Ser Thr Gly Ile Val Cys
325 330 335

Ile Ala Thr Val Arg Ser Ser Gly Lys Leu Val Ala Val Lys Lys Met
340 345 350

Asp Leu Arg Lys Gln Gln Arg Arg Glu Leu Leu Phe Asn Glu Val Val
355 360 365

Ile Met Arg Asp Tyr Gln His Glu Asn Val Val Glu Met Tyr Asn Ser
370 375 380

Tyr Leu Val Gly Asp Glu Leu Trp Val Val Met Glu Phe Leu Glu Gly
385 390 395 400

Gly Ala Leu Thr Asp Ile Val Thr His Thr Arg Met Asn Glu Glu Gln
405 410 415

Ile Ala Ala Val Cys Leu Ala Val Leu Gln Ala Leu Ser Val Leu His
420 425 430

Ala Gln Gly Val Ile His Arg Asp Ile Lys Ser Asp Ser Ile Leu Leu
435 440 445

Thr His Asp Gly Arg Val Lys Leu Ser Asp Phe Gly Phe Cys Ala Gln
450 455 460

Val Ser Lys Glu Val Pro Arg Arg Lys Ser Leu Val Gly Thr Pro Tyr
465 470 475 480

Trp Met Ala Pro Glu Leu Ile Ser Arg Leu Pro Tyr Gly Pro Glu Val
485 490 495

Asp Ile Trp Ser Leu Gly Ile Met Val Ile Glu Met Val Asp Gly Glu
500 505 510

Pro Pro Tyr Phe Asn Glu Pro Pro Leu Lys Ala Met Lys Met Ile Arg
515 520 525

Asp Asn Leu Pro Pro Arg Leu Lys Asn Leu His Lys Val Ser Pro Ser
530 535 540

Leu Lys Gly Phe Leu Asp Arg Leu Leu Val Arg Asp Pro Ala Gln Arg
 545 550 555 560

Ala Thr Ala Ala Glu Leu Leu Lys His Pro Phe Leu Ala Lys Ala Gly
 565 570 575

Pro Pro Ala Ser Ile Val Pro Leu Met Arg Gln Asn Arg Thr Arg
 580 585 590

<210> 104

<211> 3684

<212> DNA

<213> Full Length Mammalian (Human) ZC4

<400> 104

atggcgggac	ctggggcgtg	gagggacagg	gaggtcacgg	atctggcca	cctgccggat	60
ccaaactggaa	tattctca	agataaaacc	attggccttg	gtacttatgg	cagaatcttat	120
ttgggacttc	atgagaagac	tggtcattt	acagctgtta	aagtatgaa	cgctcgtaag	180
gatgaggaag	aggatctcag	gactgaactc	aacctctgtg	ggaagtactc	tttccacaaa	240
aacattgtgt	ccttctatgg	agcattttc	aagctgagtc	cccctggta	gcggcaccaa	300
cttggatgg	tgtggagtt	atgtcagca	gttccggta	ctgatgtagt	gagaatgacc	360
agtaatcaga	gttaaaaga	agattggatt	gttataatct	gccgagaaat	cttcagggc	420
ttagctcacc	ttcacgcaca	ccgagtaatt	caccgggaca	tcaaaggta	aatgtgctg	480
ctgactcata	atgctgaagt	aaaactgggt	gattttggag	tgagtgccta	ggtgaggaga	540
actaatggaa	gaaggaatag	tttcatggg	acaccatact	ggatggcacc	tgaggtgatt	600
gactgtgatg	aggacccaag	acgctcctat	gattacagaa	gtatgtgtg	gtctgtggga	660
attactgcca	ttgaaatggc	tgaaggagcc	cctcctctgt	gtaaccttca	acccttgaa	720
gctctttcg	ttattttgcg	ggaatctgt	cccacagtc	aatccagcgg	atggtcccgt	780
aagttccaca	atttcatgga	aaagtgtacg	ataaaaaatt	tcctgtttcg	tcctacttct	840
gcaaacatgc	ttcaacaccc	atttgtcgg	gatataaaaa	atgaacgaca	tgttgttgag	900
tcattaacaa	ggcatcttac	tggaatcatt	aaaaaaaaagac	agaaaaaaaga	acaggcacgg	960
gagaaaaaat	caaagtttc	tactctgagg	caagcactgg	aaaaaagact	atcaccaaag	1020
aggttcaggg	caaagtcatc	atggagacct	gaaaagcttg	aactctcgga	tttagaagcc	1080
cgcaggcaaa	ggcgcacacg	cagatggaa	gatatctta	atcagcatga	ggaagaattg	1140
agacaagttg	ataaaagacaa	agaagatgaa	tcatcagaca	atgatgaagt	atttcattcg	1200
attcaggctg	aagtccagat	agagccattt	aagccataca	tttcaaatttcc	taaaaaaaatt	1260
gaggttcaag	agagatctcc	ttctgtgcct	aacaaccagg	atcatgcaca	tcatgtcaag	1320
ttctcttcaa	gcgttcctca	gcggctcttt	ttggaaacaag	ctcagaagcc	cattgacatc	1380
agacaaagga	gttcgcacaa	tcgtcaaaat	tggctggcag	catcaggtga	ttcaaagcac	1440
aaaattttag	caggcaaaac	acagagctac	tgttaacaa	tttatatttc	agaagtcaag	1500
aaagaagaat	ttcaagaagg	aatgaatcaa	aagtgtcagg	gagcccaagt	aggatttagga	1560
cctgaaggcc	attgtatttg	gcaattgggt	gaatcttctt	ctgaggaaga	aagtcctgtg	1620
actggaaagga	ggtctcagtc	atcaccacct	tattctacta	ttgatcagaa	gttgctggtt	1680
gacatccatg	ttccagatgg	attnaaagta	ggaaaaat	caccctctgt	atacttgaca	1740
aacgaatggg	taggtataa	tgcactctt	gaaatcttcc	ggaatgattt	gttaactccg	1800
gcacctgtca	ttcagccacc	tgaagaggat	ggtgattatg	ttgaactcta	tgtgccagt	1860
gctgatactg	atggtgatga	tgtatgtag	tctaatgata	ctttgaaga	tacctatgt	1920
catgccaatg	gcaatgtatg	cttggataac	cagggtgatc	aggctaata	ttttgtaaa	1980
gaccatgtatg	atgacaacaa	taagttgtt	gatgtatgaa	ataataatta	ttatgaggcg	2040
cctagttgtc	caagggcaag	ctatggcaga	gatggaaagct	gcaagcaaga	tggttatgt	2100
ggaagtcgtg	gaaaagagga	agcctacaga	ggctatggaa	gccatacagc	caatagaagc	2160
catggaggaa	gtgcagccag	tgaggacaat	gcagccattt	gagatcagga	agaacatgca	2220
gccaatatag	gcagtgaaag	aagaggcagt	gagggtgatg	gaggttgggg	agtgcattcg	2280

accagtgaag	agagtggagc	ccttgactc	aatggagaag	aaaattgctc	agagacagat	2340
ggtccaggat	tgaagagacc	tgcgtctcg	gactttgaat	atctacagga	ggagccaggt	2400
ggtgtgaaatg	aggcctcaa	tgccattgac	tcaggtgctg	caccgtcagc	acctgatcat	2460
gagagtgaca	ataaggacat	atcagaatca	tcaacacaat	cagattttc	tgccaatcac	2520
tcatctcctt	ccaaaggttc	tgggatgtct	gctgatgcta	actttgccag	tgccatctta	2580
tacgctggat	tcgtagaagt	acctgaggaa	tcacctaagc	aaccctctga	agtcaatgtt	2640
aaccctct	atgtctctcc	tgcatgtaaa	aaaccactaa	tccacatgt	tgaaaaggag	2700
ttcacttctg	agatctgctg	tggttcttg	tggggagtca	atttgctgtt	gggaaccgga	2760
tctaatttat	atctgatgga	cagaagtgga	aaggctgaca	ttactaaact	tataaggcga	2820
agaccattcc	gccagattca	agtcttagag	ccactcaatt	tgctgattac	catctcaggt	2880
cataagaaca	gacttcgggt	gtatcatctg	acctgggtga	ggaacaagat	tttgaataat	2940
gatccagaaa	gtaaaagaag	gcaagaagaa	atgctgaaga	cagaggaagc	ctgcaaagct	3000
attgataagt	taacaggctg	tgaacacttc	agtgtcctcc	aacatgaaga	aacaacatat	3060
attgcaattg	cttgaaatc	atcaattcac	ctttatgc	gggcacccaa	gtcctttgat	3120
gaaagcaactg	ctattaaagt	atttccaaca	cttgatcata	agccagtgc	agttgacctg	3180
gctattggtt	ctgaaaaaaag	actaaagatt	ttcttcagct	cagcagatgg	atatcacctc	3240
atcgatgcag	aatctgaggt	tatgtctgat	gtgaccctgc	caaagaatcc	cctggaaatc	3300
attataccac	agaatatcat	cattttacct	gattgcttg	gaattggcat	gatgctcacc	3360
ttcaatgctg	aagccctctc	tgtggaaagca	aatgaacaaac	tcttcaagaa	gatccttgaa	3420
atgtggaaag	acataccatc	ttctatagct	tttgaatgt	cacagcgaac	cacaggatgg	3480
ggccaaaagg	ccattgaagt	gcgctctttg	caatccaggg	ttctggaaag	tgagctgaag	3540
cgcaggtcaa	ttaagaagct	gagattcctg	tgcacccggg	gtgacaagct	gttctttacc	3600
tctaccctgc	gcaatcacca	cagccgggtt	tacttcatga	cacttggaaa	acttgaagag	3660
ctccaaagca	attatgatgt	ctaa				3684

<210> 105

<211> 1227

<212> PRT

<213> Full Length Mammalian (Human) ZC4

<400> 105

Met	Ala	Gly	Pro	Gly	Gly	Trp	Arg	Asp	Arg	Glu	Val	Thr	Asp	Leu	Gly
1						5				10				15	

His	Leu	Pro	Asp	Pro	Thr	Gly	Ile	Phe	Ser	Leu	Asp	Lys	Thr	Ile	Gly
							20			25			30		

Leu	Gly	Thr	Tyr	Gly	Arg	Ile	Tyr	Leu	Gly	Leu	His	Glu	Lys	Thr	Gly
							35		40			45			

Ala	Phe	Thr	Ala	Val	Lys	Val	Met	Asn	Ala	Arg	Lys	Asp	Glu	Glu	Glu
							50		55			60			

Asp	Leu	Arg	Thr	Glu	Leu	Asn	Leu	Leu	Arg	Lys	Tyr	Ser	Phe	His	Lys
							65		70		75		80		

Asn	Ile	Val	Ser	Phe	Tyr	Gly	Ala	Phe	Phe	Lys	Leu	Ser	Pro	Pro	Gly
							85		90			95			

Gln	Arg	His	Gln	Leu	Trp	Met	Val	Met	Glu	Leu	Cys	Ala	Ala	Gly	Ser
							100		105			110			

117

Val Thr Asp Val Val Arg Met Thr Ser Asn Gln Ser Leu Lys Glu Asp
115 120 125

Trp Ile Ala Tyr Ile Cys Arg Glu Ile Leu Gln Gly Leu Ala His Leu
130 135 140

His Ala His Arg Val Ile His Arg Asp Ile Lys Gly Gln Asn Val Leu
145 150 155 160

Leu Thr His Asn Ala Glu Val Lys Leu Val Asp Phe Gly Val Ser Ala
165 170 175

Gln Val Ser Arg Thr Asn Gly Arg Arg Asn Ser Phe Ile Gly Thr Pro
180 185 190

Tyr Trp Met Ala Pro Glu Val Ile Asp Cys Asp Glu Asp Pro Arg Arg
195 200 205

Ser Tyr Asp Tyr Arg Ser Asp Val Trp Ser Val Gly Ile Thr Ala Ile
210 215 220

Glu Met Ala Glu Gly Ala Pro Pro Leu Cys Asn Leu Gln Pro Leu Glu
225 230 235 240

Ala Leu Phe Val Ile Leu Arg Glu Ser Ala Pro Thr Val Lys Ser Ser
245 250 255

Gly Trp Ser Arg Lys Phe His Asn Phe Met Glu Lys Cys Thr Ile Lys
260 265 270

Asn Phe Leu Phe Arg Pro Thr Ser Ala Asn Met Leu Gln His Pro Phe
275 280 285

Val Arg Asp Ile Lys Asn Glu Arg His Val Val Glu Ser Leu Thr Arg
290 295 300

His Leu Thr Gly Ile Ile Lys Lys Arg Gln Lys Lys Glu Gln Ala Arg
305 310 315 320

Glu Lys Lys Ser Lys Val Ser Thr Leu Arg Gln Ala Leu Ala Lys Arg
325 330 335

Leu Ser Pro Lys Arg Phe Arg Ala Lys Ser Ser Trp Arg Pro Glu Lys
340 345 350

Leu Glu Leu Ser Asp Leu Glu Ala Arg Arg Gln Arg Arg Gln Arg Arg
355 360 365

Trp Glu Asp Ile Phe Asn Gln His Glu Glu Leu Arg Gln Val Asp
370 375 380

Lys Asp Lys Glu Asp Glu Ser Ser Asp Asn Asp Glu Val Phe His Ser
385 390 395 400

118

Ile Gln Ala Glu Val Gln Ile Glu Pro Leu Lys Pro Tyr Ile Ser Asn
405 410 415

Pro Lys Lys Ile Glu Val Gln Glu Arg Ser Pro Ser Val Pro Asn Asn
420 425 430

Gln Asp His Ala His His Val Lys Phe Ser Ser Ser Val Pro Gln Arg
435 440 445

Ser Leu Leu Glu Gln Ala Gln Lys Pro Ile Asp Ile Arg Gln Arg Ser
450 455 460

Ser Gln Asn Arg Gln Asn Trp Leu Ala Ala Ser Gly Asp Ser Lys His
465 470 475 480

Lys Ile Leu Ala Gly Lys Thr Gln Ser Tyr Cys Leu Thr Ile Tyr Ile
485 490 495

Ser Glu Val Lys Lys Glu Glu Phe Gln Glu Gly Met Asn Gln Lys Cys
500 505 510

Gln Gly Ala Gln Val Gly Leu Gly Pro Glu Gly His Cys Ile Trp Gln
515 520 525

Leu Gly Glu Ser Ser Ser Glu Glu Glu Ser Pro Val Thr Gly Arg Arg
530 535 540

Ser Gln Ser Ser Pro Pro Tyr Ser Thr Ile Asp Gln Lys Leu Leu Val
545 550 555 560

Asp Ile His Val Pro Asp Gly Phe Lys Val Gly Lys Ile Ser Pro Pro
565 570 575

Val Tyr Leu Thr Asn Glu Trp Val Gly Tyr Asn Ala Leu Ser Glu Ile
580 585 590

Phe Arg Asn Asp Trp Leu Thr Pro Ala Pro Val Ile Gln Pro Pro Glu
595 600 605

Glu Asp Gly Asp Tyr Val Glu Leu Tyr Asp Ala Ser Ala Asp Thr Asp
610 615 620

Gly Asp Asp Asp Asp Glu Ser Asn Asp Thr Phe Glu Asp Thr Tyr Asp
625 630 635 640

His Ala Asn Gly Asn Asp Asp Leu Asp Asn Gln Val Asp Gln Ala Asn
645 650 655

Asp Val Cys Lys Asp His Asp Asp Asn Asn Lys Phe Val Asp Asp
660 665 670

Val Asn Asn Asn Tyr Tyr Glu Ala Pro Ser Cys Pro Arg Ala Ser Tyr
675 680 685

Gly Arg Asp Gly Ser Cys Lys Gln Asp Gly Tyr Asp Gly Ser Arg Gly
690 695 700

Lys Glu Glu Ala Tyr Arg Gly Tyr Gly Ser His Thr Ala Asn Arg Ser
705 710 715 720

His Gly Gly Ser Ala Ala Ser Glu Asp Asn Ala Ala Ile Gly Asp Gln
725 730 735

Glu Glu His Ala Ala Asn Ile Gly Ser Glu Arg Arg Gly Ser Glu Gly
740 745 750

Asp Gly Gly Lys Gly Val Val Arg Thr Ser Glu Glu Ser Gly Ala Leu
755 760 765

Gly Leu Asn Gly Glu Glu Asn Cys Ser Glu Thr Asp Gly Pro Gly Leu
770 775 780

Lys Arg Pro Ala Ser Gln Asp Phe Glu Tyr Leu Gln Glu Glu Pro Gly
785 790 795 800

Gly Gly Asn Glu Ala Ser Asn Ala Ile Asp Ser Gly Ala Ala Pro Ser
805 810 815

Ala Pro Asp His Glu Ser Asp Asn Lys Asp Ile Ser Glu Ser Ser Thr
820 825 830

Gln Ser Asp Phe Ser Ala Asn His Ser Ser Pro Ser Lys Gly Ser Gly
835 840 845

Met Ser Ala Asp Ala Asn Phe Ala Ser Ala Ile Leu Tyr Ala Gly Phe
850 855 860

Val Glu Val Pro Glu Glu Ser Pro Lys Gln Pro Ser Glu Val Asn Val
865 870 875 880

Asn Pro Leu Tyr Val Ser Pro Ala Cys Lys Lys Pro Leu Ile His Met
885 890 895

Tyr Glu Lys Glu Phe Thr Ser Glu Ile Cys Cys Gly Ser Leu Trp Gly
900 905 910

Val Asn Leu Leu Leu Gly Thr Arg Ser Asn Leu Tyr Leu Met Asp Arg
915 920 925

Ser Gly Lys Ala Asp Ile Thr Lys Leu Ile Arg Arg Pro Phe Arg
930 935 940

Gln Ile Gln Val Leu Glu Pro Leu Asn Leu Leu Ile Thr Ile Ser Gly
945 950 955 960

His Lys Asn Arg Leu Arg Val Tyr His Leu Thr Trp Leu Arg Asn Lys
965 970 975

Ile Leu Asn Asn Asp Pro Glu Ser Lys Arg Arg Gln Glu Glu Met Leu
980 985 990

120

Lys Thr Glu Glu Ala Cys Lys Ala Ile Asp Lys Leu Thr Gly Cys Glu
995 1000 1005

His Phe Ser Val Leu Gln His Glu Glu Thr Thr Tyr Ile Ala Ile Ala
1010 1015 1020

Leu Lys Ser Ser Ile His Leu Tyr Ala Trp Ala Pro Lys Ser Phe Asp
1025 1030 1035 1040

Glu Ser Thr Ala Ile Lys Val Phe Pro Thr Leu Asp His Lys Pro Val
1045 1050 1055

Thr Val Asp Leu Ala Ile Gly Ser Glu Lys Arg Leu Lys Ile Phe Phe
1060 1065 1070

Ser Ser Ala Asp Gly Tyr His Leu Ile Asp Ala Glu Ser Glu Val Met
1075 1080 1085

Ser Asp Val Thr Leu Pro Lys Asn Pro Leu Glu Ile Ile Pro Gln
1090 1095 1100

Asn Ile Ile Ile Leu Pro Asp Cys Leu Gly Ile Gly Met Met Leu Thr
1105 1110 1115 1120

Phe Asn Ala Glu Ala Leu Ser Val Glu Ala Asn Glu Gln Leu Phe Lys
1125 1130 1135

Lys Ile Leu Glu Met Trp Lys Asp Ile Pro Ser Ser Ile Ala Phe Glu
1140 1145 1150

Cys Thr Gin Arg Thr Thr Gly Trp Gly Gln Lys Ala Ile Glu Val Arg
1155 1160 1165

Ser Leu Gln Ser Arg Val Leu Glu Ser Glu Leu Lys Arg Arg Ser Ile
1170 1175 1180

Lys Lys Leu Arg Phe Leu Cys Thr Arg Gly Asp Lys Leu Phe Phe Thr
1185 1190 1195 1200

Ser Thr Leu Arg Asn His His Ser Arg Val Tyr Phe Met Thr Leu Gly
1205 1210 1215

Lys Leu Glu Glu Leu Gln Ser Asn Tyr Asp Val
1220 1225

<210> 106

<211> 2962

<212> DNA

<213> Full Length Mammalian (Human) GEK2

<400> 106

cgaagccaca	gcccggcccc	gagcccgagc	ccgagccggc	gccaccgcgc	ccccggccat	60
ggctttgcc	aatttccgcc	gcatectgcg	cctgtctacc	ttcgagaaga	gaaagtcccg	120
cgaatatgag	cacgtccgccc	gcgacctgga	ccccaaacgag	gtgtgggaga	tcgtgggca	180
gctggcgac	ggcgccctcg	gcaaggttta	caaggccaag	aataaggaga	cgggtgcctt	240
ggctgccc	aaagtcatcg	aaaccaagag	tgaggaggag	ctggaggact	acatctgtga	300
gattgagatc	ctggccaccc	gchaccaccc	ctacattgt	aagctcctgg	gagcctacta	360
tcacgacggg	aagctgtgga	tcatgattga	gttctgtcca	gggggagccg	tggacgccc	420
catgctggag	ctggacagag	gcctcacgga	gccccagata	caggtggttt	ggcccgat	480
gctagaagcc	ctcaacttcc	tgcacagcaa	gaggatcatc	caccgagatc	tgaaagctgg	540
caacgtgctg	atgaccctcg	agggagacat	caggctgct	gactttggtg	tgtctgccaa	600
aatctgaag	actctacaga	aacgagattc	tttcatcgcc	acgccttact	ggatggcccc	660
cgaggtggtc	atgtgtgaga	ccatgaaaga	cacgcctac	gactacaaag	ccgacatctg	720
gtcccggc	atcacgctga	ttgagatggc	ccagatcgag	ccggccacacc	acgagctcaa	780
ccccatgcgg	gtcctgctaa	agatgccaa	gtcggaccct	cccacgctgc	tcacgcctc	840
caagtggct	gttagagtcc	gtgacttcct	gaagatagcc	ctgataaga	acccagaaac	900
ccgaccagg	gcccgcgc	tgctggagca	tcccttcgtc	agcagcatca	ccagtaacaa	960
ggctctgcgg	gagctgggtt	ctgaggccaa	ggccgagggt	atgaaagaga	tcgaagacgg	1020
ccgggatgag	ggggaaagagg	aggacgcccgt	ggatgccc	tccaccctgg	agaaccatac	1080
tcagaactcc	tctgagggtga	gtccgccaag	cctcaatgtc	gacaagcctc	tcgaggagtc	1140
accttccacc	ccgctggcac	ccagccagtc	tcaggacagt	gtgaatgagc	cctgcagcca	1200
gccctctggg	gacagatccc	tccaaaccac	cagtccccca	gtcgtggccc	ctggaaatga	1260
gaacggcctg	gcaagtgcctg	tgccctgctg	gaagtcccg	cccggtc	tggatgccag	1320
aattcaggt	gcccaggaga	agcaagttgc	tgagcagggt	ggggaccta	gcccagcagc	1380
caacagatct	aaaaaggcca	gccagagccg	gcccaacacgc	agcgcctgg	agaccttggg	1440
tggggagaag	ctggccatg	gcagctgga	gccacctgcc	cagcagctc	cagggcctc	1500
caagaggag	tcggactgca	gcagcctctg	cacctctgag	agcatggact	atggtaccaa	1560
tctctccact	gacctgtcgc	tgaacaaaga	gatgggctt	ctgtccatca	aggacccgaa	1620
actgtacaaa	aaaaccctca	agcggacacg	caaatttgg	gtggatggtg	tggaggtgag	1680
catcaccacc	tccaaagatca	tcagcgaaga	tgagaagaag	gatgaggaga	tgagatttct	1740
caggcgc	gaactccgag	agcttcggct	gtccagaaa	gaagagcatc	ggaaccagac	1800
ccagctgagt	aacaagcatg	agctgcagct	ggagcaa	cataaacgtt	ttgaacagga	1860
aatcaacgc	aagaagaagt	tcttgacac	gaaatttagag	aacctggagc	gtca	1920
gcagcaagt	gagaagatgg	agcaagacca	tgccgtgcgc	cgccggagg	aggccaggcg	1980
gatccgcctg	gagcaggatc	gggactacac	caggtccaa	gagcagctca	aactgtgaa	2040
gaaagaggt	aagaacgagg	tggagaagct	cccccgacag	cagcggagg	aaagcatgaa	2100
gcagaagatg	gaggagcaca	cgcaaaaaaa	gagcttctt	gaccggact	tttgtagccaa	2160
gcagaaggag	gac	ttggccatgaa	gaggtcacc	accgacaaca	ggcgggagat	2220
ctgtgacaag	gagcgcgagt	gcctcatgaa	gaagcaggag	ctccttcgag	accggaa	2280
agccctgtgg	gagatggaa	agcaccagct	gcaggagagg	caccagctgg	tgaagcagca	2340
gctcaaagac	cagacttcc	tccagcggca	cgagctgctg	cgcaagcatg	agaaggagcg	2400
ggagcagatg	cagcgc	accagcgcat	gatagagcag	ctgaagggtgc	ggcagcaaca	2460
ggaaaaggcg	cggtgtccca	agatccagag	gagtggggc	aagacgcgc	tggccatgta	2520
caagaagagc	ctccacatca	acggcgggg	cagcgcagct	gagcagcgt	agaagatcaa	2580
gcagttctcc	cagcaggagg	agaagaggca	gaagtccggag	cggtgcgc	aacagcagaa	2640
acacgagaac	cagatgcggg	acatgtggc	gcagtgcgag	agcaacatga	gcgagctgca	2700
gcagctgcag	aatgaaaagt	gccacccct	gttagagcac	gaaacccaga	aactgaaggc	2760
cctggatgag	agccataacc	agaacctgaa	gaaatggcgg	gacaagctt	ggccgcgca	2820
gaaggctctg	gaagaggatc	tgaaccagaa	gaagcggggag	caggagatgt	tcttcaagct	2880
gagcggaggag	gcggagtgcc	caaaccctc	cacccaaagc	aaggccgcca	agttcttccc	2940
ctacagctct	ggggatgctt	cc				2962

<210> 107
<211> 968
<212> PRT
<213> Full Length Mammalian (Human) GEK2

<400> 107

Met Ala Phe Ala Asn Phe Arg Arg Ile Leu Arg Leu Ser Thr Phe Glu
1 5 10 15

Lys Arg Lys Ser Arg Glu Tyr Glu His Val Arg Arg Asp Leu Asp Pro
20 25 30

Asn Glu Val Trp Glu Ile Val Gly Glu Leu Gly Asp Gly Ala Phe Gly
35 40 45

Lys Val Tyr Lys Ala Lys Asn Lys Glu Thr Gly Ala Leu Ala Ala Ala
50 55 60

Lys Val Ile Glu Thr Lys Ser Glu Glu Glu Leu Glu Asp Tyr Ile Val
65 70 75 80

Glu Ile Glu Ile Leu Ala Thr Cys Asp His Pro Tyr Ile Val Lys Leu
85 90 95

Leu Gly Ala Tyr Tyr His Asp Gly Lys Leu Trp Ile Met Ile Glu Phe
100 105 110

Cys Pro Gly Gly Ala Val Asp Ala Ile Met Leu Glu Leu Asp Arg Gly
115 120 125

Leu Thr Glu Pro Gln Ile Gln Val Val Cys Arg Gln Met Leu Glu Ala
130 135 140

Leu Asn Phe Leu His Ser Lys Arg Ile Ile His Arg Asp Leu Lys Ala
145 150 155 160

Gly Asn Val Leu Met Thr Leu Glu Gly Asp Ile Arg Leu Ala Asp Phe
165 170 175

Gly Val Ser Ala Lys Asn Leu Lys Thr Leu Gln Lys Arg Asp Ser Phe
180 185 190

Ile Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Val Met Cys Glu Thr
195 200 205

Met Lys Asp Thr Pro Tyr Asp Tyr Lys Ala Asp Ile Trp Ser Leu Gly
210 215 220

Ile Thr Leu Ile Glu Met Ala Gln Ile Glu Pro Pro His His Glu Leu
225 230 235 240

Asn Pro Met Arg Val Leu Leu Lys Ile Ala Lys Ser Asp Pro Pro Thr
245 250 255

123

Leu Leu Thr Pro Ser Lys Trp Ser Val Glu Phe Arg Asp Phe Leu Lys
260 265 270

Ile Ala Leu Asp Lys Asn Pro Glu Thr Arg Pro Ser Ala Ala Gln Leu
275 280 285

Leu Glu His Pro Phe Val Ser Ser Ile Thr Ser Asn Lys Ala Leu Arg
290 295 300

Glu Leu Val Ala Glu Ala Lys Ala Glu Val Met Glu Glu Ile Glu Asp
305 310 315 320

Gly Arg Asp Glu Gly Glu Glu Asp Ala Val Asp Ala Ala Ser Thr
325 330 335

Leu Glu Asn His Thr Gln Asn Ser Ser Glu Val Ser Pro Pro Ser Leu
340 345 350

Asn Ala Asp Lys Pro Leu Glu Glu Ser Pro Ser Thr Pro Leu Ala Pro
355 360 365

Ser Gln Ser Gln Asp Ser Val Asn Glu Pro Cys Ser Gln Pro Ser Gly
370 375 380

Asp Arg Ser Leu Gln Thr Thr Ser Pro Pro Val Val Ala Pro Gly Asn
385 390 395 400

Glu Asn Gly Leu Ala Val Pro Val Pro Leu Arg Lys Ser Arg Pro Val
405 410 415

Ser Met Asp Ala Arg Ile Gln Val Ala Gln Glu Lys Gln Val Ala Glu
420 425 430

Gln Gly Gly Asp Leu Ser Pro Ala Ala Asn Arg Ser Gln Lys Ala Ser
435 440 445

Gln Ser Arg Pro Asn Ser Ser Ala Leu Glu Thr Leu Gly Gly Glu Lys
450 455 460

Leu Ala Asn Gly Ser Leu Glu Pro Pro Ala Gln Ala Ala Pro Gly Pro
465 470 475 480

Ser Lys Arg Asp Ser Asp Cys Ser Ser Leu Cys Thr Ser Glu Ser Met
485 490 495

Asp Tyr Gly Thr Asn Leu Ser Thr Asp Leu Ser Leu Asn Lys Glu Met
500 505 510

Gly Ser Leu Ser Ile Lys Asp Pro Lys Leu Tyr Lys Lys Thr Leu Lys
515 520 525

Arg Thr Arg Lys Phe Val Val Asp Gly Val Glu Val Ser Ile Thr Thr
530 535 540

Ser Lys Ile Ile Ser Glu Asp Glu Lys Lys Asp Glu Glu Met Arg Phe
545 550 555 560

Leu Arg Arg Gln Glu Leu Arg Glu Leu Arg Leu Leu Gln Lys Glu Glu
565 570 575

His Arg Asn Gln Thr Gln Leu Ser Asn Lys His Glu Leu Gln Leu Glu
580 585 590

Gln Met His Lys Arg Phe Glu Gln Glu Ile Asn Ala Lys Lys Lys Phe
595 600 605

Phe Asp Thr Glu Leu Glu Asn Leu Glu Arg Gln Gln Lys Gln Gln Val
610 615 620

Glu Lys Met Glu Gln Asp His Ala Val Arg Arg Glu Glu Ala Arg
625 630 635 640

Arg Ile Arg Leu Glu Gln Asp Arg Asp Tyr Thr Arg Phe Gln Glu Gln
645 650 655

Leu Lys Leu Met Lys Lys Glu Val Lys Asn Glu Val Glu Lys Leu Pro
660 665 670

Arg Gln Gln Arg Lys Glu Ser Met Lys Gln Lys Met Glu Glu His Thr
675 680 685

Gln Lys Lys Gln Leu Leu Asp Arg Asp Phe Val Ala Lys Gln Lys Glu
690 695 700

Asp Leu Glu Leu Ala Met Lys Arg Leu Thr Thr Asp Asn Arg Arg Glu
705 710 715 720

Ile Cys Asp Lys Glu Arg Glu Cys Leu Met Lys Lys Gln Glu Leu Leu
725 730 735

Arg Asp Arg Glu Ala Ala Leu Trp Glu Met Glu Glu His Gln Leu Gln
740 745 750

Glu Arg His Gln Leu Val Lys Gln Gln Leu Lys Asp Gln Tyr Phe Leu
755 760 765

Gln Arg His Glu Leu Leu Arg Lys His Glu Lys Glu Arg Glu Gln Met
770 775 780 785

Gln Arg Tyr Asn Gln Arg Met Ile Glu Gln Leu Lys Val Arg Gln Gln
790 795 800

Gln Glu Lys Ala Arg Leu Pro Lys Ile Gln Arg Ser Glu Gly Lys Thr
805 810 815

Arg Met Ala Met Tyr Lys Ser Leu His Ile Asn Gly Gly Gly Ser
820 825 830

Ala Ala Glu Gln Arg Glu Lys Ile Lys Gln Phe Ser Gln Gln Glu Glu
835 840 845

125

Lys Arg Gln Lys Ser Glu Arg Leu Gln Gln Gln Lys His Glu Asn
850 855 860

Gln Met Arg Asp Met Leu Ala Gln Cys Glu Ser Asn Met Ser Glu Leu
865 870 875 880

Gln Gln Leu Gln Asn Glu Lys Cys His Leu Leu Val Glu His Glu Thr
885 890 895

Gln Lys Leu Lys Ala Leu Asp Glu Ser His Asn Gln Asn Leu Lys Glu
900 905 910

Trp Arg Asp Lys Leu Arg Pro Arg Lys Lys Ala Leu Glu Glu Asp Leu
915 920 925

Asn Gln Lys Lys Arg Glu Gln Glu Met Phe Phe Lys Leu Ser Glu Glu
930 935 940

Ala Glu Cys Pro Asn Pro Ser Thr Pro Ser Lys Ala Ala Lys Phe Phe
945 950 955 960

Pro Tyr Ser Ser Gly Asp Ala Ser
965

<210> 108

<211> 11

<212> PRT

<213> Mammalian (human) ZC1 and/or ZC3

<220>

<223> "Xaa can stand for any amino acid

<400> 108

Leu Xaa Leu Xaa Leu Xaa Leu Xaa Leu Xaa Leu
1 5 10

<210> 109

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized nucleic acid molecule

<400> 109

gcagcaagtg gagaagatgg

20

<210> 110

126

<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized nucleic acid molecule

<400> 110

ggaaggcatcc ccagagctgt ag

22

<210> 111
<211> 12
<212> PRT
<213> Mammalian (Human) STLK2

<400> 111

Glu Lys Phe Gln Lys Cys Ser Ala Asp Glu Ser Pro
1 5 10

<210> 112
<211> 16
<212> PRT
<213> Mammalian (Human) STLK4

<400> 112

Ser Ile Ser Asn Ser Glu Leu Phe Pro Thr Thr Asp Pro Val Gly Thr
1 5 10 15

<210> 113
<211> 9
<212> PRT
<213> Mammalian (Human) SULU1

<400> 113

Leu Asp Phe Pro Lys Glu Asp Tyr Arg
1 5

<210> 114
<211> 12
<212> PRT
<213> Mammalian (Human) SULU1

<400> 114

His Gly Asp Pro Arg Pro Glu Pro Arg Pro Thr Gln
1 5 10

<210> 115
<211> 14
<212> PRT
<213> Mammalian (Human) SULU3

<400> 115

Pro Ser Thr Asn Arg Ala Gly Ser Leu Lys Asp Pro Glu Cys
1 5 10

<210> 116
<211> 19
<212> PRT
<213> Mammalian (Human) SULU3

<400> 116

Asp Pro Arg Thr Arg Ala Ser Asp Pro Gln Ser Pro Pro Gln Val Ser
1 5 10 15

Arg His Lys

<210> 117
<211> 16
<212> PRT
<213> Mammalian (Human) PAK4

<400> 117

Cys Leu Val Pro Leu Ile Gln Leu Tyr Arg Lys Gln Thr Ser Thr Cys
1 5 10 15

<210> 118
<211> 9
<212> PRT
<213> Mammalian (Human) PAK5

<400> 118

Pro Leu Met Arg Gln Asn Arg Thr Arg
1 5

<210> 119
<211> 16
<212> PRT
<213> Mammalian (Human) PAK5

<400> 119

Ser Gly Asp Arg Arg Arg Ala Gly Pro Glu Lys Arg Pro Lys Ser Ser
1 5 10 15

<210> 120
<211> 16
<212> PRT
<213> Mammalian (Human) PAK5

<400> 120

Cys Arg Arg Lys Ser Leu Val Gly Thr Pro Tyr Trp Met Ala Pro Glu
1 5 10 15

<210> 121
<211> 20
<212> PRT
<213> Mammalian (Human) ZC1

<400> 121

Arg Thr Val Gly Arg Arg Asn Thr Phe Ile Gly Thr Pro Pro Tyr Trp
1 5 10 15

Met Ala Pro Glu
20

<210> 122
<211> 15
<212> PRT
<213> Mammalian (Human) KHS1

<400> 122

Lys Arg Lys Ser Phe Ile Gly Thr Pro Tyr Trp Met Ala Pro Glu
1 5 10 15

<210> 123
<211> 15
<212> PRT
<213> Mammalian (Human) STLK2

<400> 123

Lys Arg Asn Thr Phe Val Gly Thr Pro Phe Trp Met Ala Pro Glu
1 5 10 15

<210> 124
<211> 15
<212> PRT
<213> Mammalian (Human) SULU1

<400> 124

Pro Ala Asn Ser Phe Val Gly Thr Pro Tyr Trp Met Ala Pro Glu
1 5 10 15

<210> 125
<211> 15
<212> PRT
<213> Mammalian (Human) ZC1

<400> 125

Arg Arg Asn Thr Phe Ile Gly Thr Pro Tyr Trp Met Ala Pro Glu
1 5 10 15

<210> 126
<211> 15
<212> PRT
<213> Mammalian (Human) ZC1

<400> 126

Arg Arg Asn Thr Phe Ile Gly Thr Pro Tyr Trp Met Ala Pro Glu
1 5 10 15

<210> 127
<211> 18
<212> PRT
<213> Mammalian (Human) STLK4

130

<400> 127

Arg Asn Lys Val Arg Lys Thr Phe Val Gly Thr Pro Cys Trp Met Ala
1 5 10 15

Pro Glu

<210> 128

<211> 22

<212> DNA

<213> Mammalian (Human) STLK3

<400> 128

ctccccatttc ctagcaaaat ca

22

<210> 129

<211> 22

<212> DNA

<213> Mammalian (Human) STLK3

<400> 129

agaggcagta ttgtcagatg ta

22

<210> 130

<211> 22

<212> DNA

<213> Mammalian (Human) STLK4

<400> 130

ccacacatgc gatatctgt tg

22

<210> 131

<211> 24

<212> DNA

<213> Mammalian (Human) STLK4

<400>

131

ttgctagaat tcacatcagg taca

24

<210> 132
<211> 22
<212> DNA
<213> Mammalian (Human) ZC1

<400> 132

atccctggat cacactgctt ct

22

<210> 133
<211> 22
<212> DNA
<213> Mammalian (Human) ZC1

<400> 133

caaggtgttc tttgcctctg tt

22

<210> 134
<211> 21
<212> DNA
<213> Mammalian (Human) ZC2

<400> 134

agatggactg tactgggagg g

21

<210> 135
<211> 22
<212> DNA
<213> Mammalian (Human) ZC2

<400> 135

agaagagcac ttggcactta tc

22

<210> 136
<211> 20
<212> DNA
<213> Mammalian (Human) ZC3

132

<400> 136

catcatgaac tggtgacggg

20

<210> 137

<211> 23

<212> DNA

<213> Mammalian (Human) ZC3

<400> 137

ccagtgaat caaaccagta aaa

23

<210> 138

<211> 24

<212> DNA

<213> Mammalian (Human) SULU1

<400> 138

caaaaacctgg ccgtctcttc tatt

24

<210> 139

<211> 23

<212> DNA

<213> Mammalian (Human) SULU1

<400> 139

attttgtgcta ctgggattct gtg

23

<210> 140

<211> 24

<212> DNA

<213> Mammalian (Human) KHS2

<400> 140

gaatagcggt accatgatag aata

24

<210> 141

133

<211> 23
<212> DNA
<213> Mammalian (Human) KHS2

<400> 141

tacccaaaaag agccaaaagt gtg

23

<210> 142
<211> 23
<212> DNA
<213> Mammalian (Human) PAK4
<400> 142

ctcagttattc tctccaaaga ttg

23

<210> 143
<211> 23
<212> DNA
<213> Mammalian (Human) PAK4
<400> 143

gatgttctct ccattctgtta aag

23

<210> 144
<211> 21
<212> DNA
<213> Mammalian (Human) PAK5
<400> 144

catcaactgga agtctgcagt g

21

<210> 145
<211> 20
<212> DNA
<213> Mammalian (Human) PAK5
<400> 145

cagggtgcagt agtcatttg

20

<210> 146
<211> 20
<212> DNA
<213> T7 phage

<400> 146

ggagctgtcg tattccagtc

20

<210> 147
<211> 21
<212> DNA
<213> T7 phage

<400> 147

aaccctcaa gacccgttta g

21



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 9/12, 15/52		A3	(11) International Publication Number: WO 99/53036
			(43) International Publication Date: 21 October 1999 (21.10.99)
(21) International Application Number: PCT/US99/08150		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 13 April 1999 (13.04.99)			
(30) Priority Data: 60/081,784 14 April 1998 (14.04.98) US			
(71) Applicant (for all designated States except US): SUGEN, INC. [US/US]; 230 East Grand Avenue, South San Francisco, CA 94080 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): PLOWMAN, Gregory [US/US]; 4 Honeysuckle Lane, San Carlos, CA 94070 (US). MARTINEZ, Ricardo [US/US]; 984 Cartier Lane, Foster City, CA 94404 (US). WHYTE, David [US/US]; 2623 Barclay Way, Belmont, CA 94002 (US).			
(74) Agents: WARBURG, Richard, J. et al.; Lyon & Lyon LLP, Suite 4700, 633 West Fifth Street, Los Angeles, CA 90071-2066 (US).			
<p>Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>			
(88) Date of publication of the international search report: 11 May 2000 (11.05.00)			

(54) Title: STE20-RELATED PROTEIN KINASES

(57) Abstract

The present invention relates to the novel kinase polypeptides STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEK2, PAK4, and PAK5, nucleotide sequences encoding the novel kinase polypeptides, as well as various products and methods useful for the diagnosis and treatment of various kinase-related diseases and conditions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/08150

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 6 C12N9/12 C12N15/52

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched.

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>WO 97 42212 A (GEN HOSPITAL CORP) 13 November 1997 (1997-11-13)</p> <p>page 2, line 18 -page 9, line 12 page 10, line 9 -page 11, line 23 page 17, line 11 -page 19, line 13 All inventions</p> <p>---</p> <p>BUCHER ET AL: "A flexible motif search technique based on generalized profiles" COMPUTERS AND CHEMISTRY, GB, PERGAMON PRESS, OXFORD, vol. 20, no. 1, 1996, pages 3-23, XP002107535 ISSN: 0097-8485 the whole document All inventions</p> <p>---</p>	1-15, 21-24, 29-32
Y		1-15, 21-24, 29-32

Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

8 March 2000

Date of mailing of the international search report

17.03.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Sprinks, M

INTERNATIONAL SEARCH REPORT

Intern'l Application No
PCT/US 99/08150

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		Relevant to claim No.
Category	Citation of document, with indication, where appropriate, of the relevant passages	
P,X	WO 99 15635 A (ZENECA LTD) 1 April 1999 (1999-04-01) page 6, line 16 -page 8, line 13; claims 1-19; figures 1-10 Invention 1 ---	1-15, 21-24, 29-32
E	WO 99 32637 A (ZENECA LTD) 1 July 1999 (1999-07-01) page 7, line 3 -page 9, line 24; claims 1-19; figures 1-16 Invention 1 ---	1-15, 21-24, 29-32
P,X	WO 99 07854 A (MIAO NINGNING ;ONTOGENY INC (US); PANG KEVIN (US); BARKER DOUGLAS) 18 February 1999 (1999-02-18) claims 1-44; table 1 Invention 2 - see SEQ ID NO:5/6 Invention 3 - see SEQ ID NO:7/8/9 Invention 6 - see SEQ ID NO:7/8/9 ---	1-15, 21-24, 29-32 21-24, 29-32
P,Y	DATABASE EMBL [Online] ID: AF099989, 11 November 1998 (1998-11-11) JOHNSTON ET AL.: "SPAK: a novel Ste-20 related kinase expressed in the pancreas" XP002132350 abstract Invention 2 ---	1-15
P,Y	Invention 2 ---	21-24, 29-32
X	DATABASE EMBL [Online] ID: AF017635, 23 September 1997 (1997-09-23) BAYTEL ET AL.: "Homo sapiens DCHT mRNA, complete cds" XP002132351 abstract Invention 2 - clearly encodes a kinase ---	1-15
Y	DATABASE EMBL [Online] ID: MMAA20708, 21 November 1996 (1996-11-21) MARRA ET AL.: "mp54a01.r1 Soares 2NbMT Mus musculus cDNA clone 573000 5'" XP002132352 abstract Inventions 4 and 5 - no function indicated ---	21-24, 29-32
X	-/-	1-9

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/US 99/08150

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL [Online] ID: HS130B11B, 25 August 1995 (1995-08-25) FUJIWARA ET AL.: "Human fetal brain cDNA 5'-end GEN-130B11" XP002132353 abstract Invention 4 - no function indicated ---</p>	1-9
X	<p>DATABASE EMBL [Online] ID: AA766905, 30 January 1998 (1998-01-30) NCI-CGAP: "...Homo sapiens cDNA clone Image:1301771 similar to TR:Q42341 Q42341 SERINE-THREONINE PROTEIN KINASE" XP002132354 abstract</p>	1-15
Y	<p>Invention 6 - encoded function indicated ---</p>	21-24, 29-32
X	<p>SU ET AL.: "NIK is a new Ste20-related kinase that binds NCK and MEKK1 and activates the SAPK/JNK cascade via a conserved regulatory domain" THE EMBO JOURNAL, vol. 16, no. 6, 1997, pages 1279-1290, XP002132378 abstract; figure 1</p>	1-14,16, 21-24
Y	<p>Invention 7 - identical to residues 1-47, essentially identical from 1-495 and 625-1239 (end)</p>	29-32
X	<p>DATABASE EMBL [Online] ID: AB011123, 10 April 1998 (1998-04-10) OHARA ET AL.: "Homo sapiens mRNA for KIAA0551 protein, partial cds" XP002132377 abstract</p>	1-14,16
Y	<p>Invention 8 - almost 100% identical to residues 8-410 and 415-1297 (end) Invention 9 - shows significant sequence similarity ---</p>	21-24, 29-32
X	<p>DATABASE EMBL [Online] ID: AA865818, 16 March 1998 (1998-03-16) NCI-CGAP: "...Homo sapiens cDNA clone IMAGE:1456752 3' similar to TR:P97820 P97820 NIK..." XP002132508 abstract</p>	1-14,16
Y	<p>Invention 10 - encoded function indicated (similar to NIK, a known protein kinase) --- -/-</p>	21-24, 29-32

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/US 99/08150

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		Relevant to claim No.
Category *	Citation of document, with indication, where appropriate, of the relevant passages	
X	DATABASE EMBL [Online] ID: HS571200, 15 September 1995 (1995-09-15) HILLIER ET AL.: "yr32h11.r1 Homo sapiens cDNA clone 207045 5'" XP002132509 abstract Invention 10 - no function indicated ---	1-9
X	DATABASE EMBL [Online] ID: HS1254577, 13 June 1997 (1997-06-13) HILLIER ET AL.: "...Homo sapiens cDNA clone 796310 5' similar to WP:ZC504.4 CE02384 SERINE/THREONINE PROTEIN KINASE" XP002132510 abstract	1-14, 16
Y	Invention 10 - encoded function indicated ---	21-24, 29-32
X	DIENER ET AL.: "Activation of the c.Jun N-terminal kinase pathway by a novel protein kinase related to human germinal center kinase" PROC. NATL. ACAD. SCI. USA, vol. 94, September 1997 (1997-09), pages 9687-9692, XP002132504 abstract; figures 1-7	1-14, 17, 21-24
Y	Invention 11 - GLK is identical to residues 13-391 and 393-894 (end) of KHS2 ---	29-32
P,X	WO 99 02699 A (CADUS PHARMACEUTICAL CORP) 21 January 1999 (1999-01-21) abstract; figure 2 Invention 12 ---	1-14, 18, 21-24, 29-32
X	DATABASE EMBL [Online] ID: AA885355, 30 March 1998 (1998-03-30) NCI-CGAP: "...Homo sapiens cDNA clone IMAGE:1460315 3' similar to WP:T17E9.1 CE01405" XP002132511 abstract Invention 12 - no function indicated ---	1-9
X	DATABASE EMBL [Online] ID: AA576724, 11 September 1997 (1997-09-11) NCI-CGAP: "...Homo sapiens cDNA clone IMAGE:1074607" XP002132512 abstract Invention 12 - no function indicated ---	1-9

INTERNATIONAL SEARCH REPORT

Intern	al Application No
PCT/US 99/08150	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] ID: MM1266197, 22 June 1997 (1997-06-22) MARRA ET AL.: "...Mus musculus cDNA clone 805425 5' similar to WP:T17E9.1 CE01405" XP002132513 abstract Invention 13 - no function indicated ---	1-9
P,X	HUTCHISON M: "Isolation of TA01, a protein kinase that activates MEKs in stress-activated protein kinase cascades" JOURNAL OF BIOLOGICAL CHEMISTRY, US, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, vol. 273, no. 44, 30 October 1998 (1998-10-30), pages 28625-28632-28632, XP002114118 ISSN: 0021-9258 abstract; figures 1-6 Invention 13 - TA01 is essentially identical to SULU3 ---	1-14, 18, 21-24
Y	KURAMOCHI ET AL.: "LOK is a novel mouse STE20-like protein kinase that is expressed predominantly in lymphocytes" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 272, no. 36, 5 September 1997 (1997-09-05), pages 22679-22684, XP002132505 abstract; figures 1-6 Invention 14 - LOK is essentially identical to (probably the mouse homologue of) GEK2 - identical at amino acid positions 1-33 ---	29-32
X	DATABASE EMBL [Online] ID: HS1259479, 20 June 1997 (1997-06-20) NCI-CGAP: "...Homo sapiens cDNA clone IMAGE:814858 3' similar to TR:G881958 G881958 MESS1" XP002132515 abstract Invention 14 - encoded function indicated (MESS1 is a protein kinase of the prior art) ---	1-14, 19
Y	---	21-24, 29-32
		-/-

INTERNATIONAL SEARCH REPORT

Intern: Application No:
PCT/US 99/08150

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] ID: HS1254308, 16 June 1997 (1997-06-16) NCI-CGAP: "...Homo sapiens cDNA clone IMAGE:814858 5' similar to WP:T19A5.2 CE07510 SERINE.THREONINE KINASE" XP002132516 abstract	1-14, 19
Y	Invention 14 - encoded function indicated ---	21-24, 29-32
P,X	DATABASE EMBL [Online] ID: AB015718, 14 December 1998 (1998-12-14) KURAMOCHI ET AL.: "Homo sapiens 1ok mRNA for protein kinase, complete cds" XP002132514 abstract	1-14, 19, 21-24
P,Y	Invention 14 - STK10 is identical to GEK2 at all but one amino acid positions ---	29-32
X	DATABASE EMBL [Online] ID: AA634299, 31 October 1997 (1997-10-31) HILLIER ET AL.: "...Homo sapiens cDNA clone 743770 3'" XP002132517 abstract	1-9
P,X	Invention 15 - no function indicated ---	1-14, 20-24
P,Y A	ABO ET AL.: "PAK4, a novel effector for Cdc42Hs, is implicated in the reorganization of the actin cytoskeleton and in the formation of filopodia" THE EMBO JOURNAL, vol. 17, no. 22, 16 November 1998 (1998-11-16), pages 6527-6540, XP002132507 abstract; figures 1-8 "A" for invention 15 - the name "PAK4" has been given to different proteins "PX", "PY" and "A" for invention 16 - "PAK5" of the present application is identical to "PAK4" of this document. -----	29-32 1-14, 20-24, 29-32

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/08150

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

2. 25-28 Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 25-28

It is not possible to carry out a meaningful search into the state of the art on the basis of claims 25-28 because they refer to the use of "modulators" and "kinase inhibitors" which are structurally undefined and could not in any event have been functionally tested in the prior art (assuming novelty for the kinases to which they refer).

The applicant is also requested to note that additional problems during subsequent examination may also result from the formulation of said claims, which currently refer to methods of treatment of the human or animal body.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-15,21-24,29-32 partially

A nucleic acid encoding a STLK2 kinase polypeptide (SEQ ID NO:5) and subject-matter relating thereto.

2. Claims: 1-15,21-24,29-32 partially

A nucleic acid encoding a STLK3 kinase polypeptide (SEQ ID NO:6) and subject-matter relating thereto.

3. Claims: 1-15,21-24,29-32 partially

A nucleic acid encoding a STLK4 kinase polypeptide (SEQ ID NO:7) and subject-matter relating thereto.

4. Claims: 1-15,21-24,29-32 partially

A nucleic acid encoding a STLK5 kinase polypeptide (SEQ ID NO:97) and subject-matter relating thereto.

5. Claims: 1-15,21-24,29-32 partially

A nucleic acid encoding a STLK6 kinase polypeptide (SEQ ID NO:99) and subject-matter relating thereto.

6. Claims: 1-15,21-24,29-32 partially

A nucleic acid encoding a STLK7 kinase polypeptide (SEQ ID NO:101) and subject-matter relating thereto.

7. Claims: 1-14,16,21-24,29-32 partially

A nucleic acid encoding a ZC1 kinase polypeptide (SEQ ID NO:13) and subject-matter relating thereto.

8. Claims: 1-14,16,21-24,29-32 partially

A nucleic acid encoding a ZC2 kinase polypeptide (SEQ ID NO:14) and subject-matter relating thereto.

9. Claims: 1-14,16,21-24,29-32 partially

A nucleic acid encoding a ZC3 kinase polypeptide (SEQ ID NO:15) and subject-matter relating thereto.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10. Claims: 1-14,16,21-24,29-32 partially

A nucleic acid encoding a ZC4 kinase polypeptide (SEQ ID NO:105) and subject-matter relating thereto.

11. Claims: 1-14,21-24,29-32 partially; 17 completely

A nucleic acid encoding a KHS2 kinase polypeptide (SEQ ID NO:18) and subject-matter relating thereto.

12. Claims: 1-14,18,21-24,29-32 partially

A nucleic acid encoding a SULU1 kinase polypeptide (SEQ ID NO:22) and subject-matter relating thereto.

13. Claims: 1-14,18,21-24,29-32 partially

A nucleic acid encoding a SULU3 kinase polypeptide (SEQ ID NO:23) and subject-matter relating thereto.

14. Claims: 1-14,21-24,29-32 partially; 19 completely

A nucleic acid encoding a GEK2 kinase polypeptide (SEQ ID NO:107) and subject-matter relating thereto.

15. Claims: 1-14,20-24,29-32 partially

A nucleic acid encoding a PAK4 kinase polypeptide (SEQ ID NO:29) and subject-matter relating thereto.

16. Claims: 1-14,20-24,29-32 partially

A nucleic acid encoding a PAK5 kinase polypeptide (SEQ ID NO:103) and subject-matter relating thereto.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Internal Application No

PCT/US 99/08150

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9742212	A 13-11-1997	AU 3118297 A US 5830699 A		26-11-1997 03-11-1998
WO 9915635	A 01-04-1999	AU 9172698 A		12-04-1999
WO 9932637	A 01-07-1999	AU 1676699 A US 5962265 A		12-07-1999 05-10-1999
WO 9907854	A 18-02-1999	AU 8778698 A		01-03-1999
WO 9902699	A 21-01-1999	AU 8296698 A		08-02-1999

THIS PAGE BLANK (USPTO)